DOCUMENT RESUME

ED 113 013 PS 008 002

AUTHOR Lukas, Carol VanDeusen; Wohlleb, Cindy
TITLE Implementation in Head Start Planned Variation: 1971-72.
INSTITUTION Huron Inst., Cambridge, Mass.
SPONS AGENCY Office of Child Development (DHEW), Washington, D.C.
REPORT NO OCD-H-1926
PUB DATE Jun 74
NOTE 97p.; For related documents, see PS 008 000-001 and ED 082 834-835; Sponsor's evaluation form not included due to poor reproducibility

EDRS PRICE MF-$0.76 HC-$4.43 Plus Postage
DESCRIPTORS Check Lists; Classroom Observation Techniques; *Comparative Analysis; *Compensatory Education Programs; Data Analysis; *Early Childhood Education; Longitudinal Studies; Objective Tests; Program Effectiveness; *Program Evaluation; *Teacher Evaluation
IDENTIFIERS *Project Head Start

ABSTRACT

This report on model implementation during the third year (1971-72) of the Head Start Planned Variation experiment concentrates on the development of measures of implementation. Measures are examined individually in terms of design and use, strengths and weaknesses, and findings obtained, and agreement among raters on the extent of implementation is discussed. A comparative analysis of consultant checklists, sponsor checklists, sponsor overall ratings, teacher perceptions, and sponsor ratings is presented. Proposals to improve implementation measures include a contractual agreement between evaluators and those responsible for treatment implementation, field testing of implementation measures prior to their use, and systematic training of raters. The report concludes that substantial problems remain with implementation measurement, that some models may never fit into a quantified, easily observable measure, and that there may always be an element of subjectivity in assessing the extent to which a program is implemented. Appendices comprise site to site variation in levels of implementation and the 1971-72 implementation measures. (GO)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED113013

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

IMPLEMENTATION IN HEAD START

PLANNED VARIATION: 1971-72

Carol VanDeusen Lukas
Cindy Wohlleb

The Huron Institute
Cambridge, Massachusetts
June, 1974

This document was prepared for Grant # H 1926 from the Office of Child Development, Department of Health, Education and Welfare, U. S. Government. The conclusions and recommendations in this report are those of the grantee and do not necessarily reflect the views of any federal agency.

The project directors were Marshall S. Smith & Herbert I. Weisberg

PS 008002

00002

Table of Contents

	Page
Issues of Implementation	1
The 1970-71 Implementation Study	4
The 1971-72 Analysis	7
Measures of Implementation	11
Consultant Checklists	11
Sponsor Checklists	20
Sponsor Overall Ratings	23
Teacher Perceptions	25
Sponsor Ratings	27
Summary	33
Correspondence Among Implementation Measures	35
Conclusions	44
Appendices	

00063

ISSUES OF IMPLEMENTATION

Head Start Planned Variation (HSPV) was a federally funded experiment designed to compare the effectiveness of a variety of methods of preschool education. The study involved the installation of each of eleven different programs, or models, in one to five sites across the country. The sites were Head Start centers which were already operating at the beginning of the study. In most cases, not all of the classrooms in a center were designated as Planned Variation classes; some classrooms continued as regular (or comparison) Head Start classes. Table 1 presents information about the models and Head Start centers participating in the Planned Variation experiment.

The responsibility for putting the model theories into practice in the sites was given to the model sponsors, the individuals or organizations who designed the programs. Sponsors provided program literature and materials, introduced the model to the community, and trained the staff to use the model's teaching methods and curriculum.

Although the primary focus of HSPV was on testing the effects of these models on children, the purpose of this report is to study the implementation of the models. In studying implementation we are interested both in the implementation process -- the ways in which a model is transferred to a site and the factors

TABLE 1

Head Start Planned Variation Sites; 1969-72

	1969-70	1970-71	1971-72
Far West	Duluth	Duluth Fresno Tacoma Salt Lake Buffalo	Duluth Salt Lake Tacoma Buffalo
Arizona	LaFayette Lakewood	LaFayette Lakewood Lincoln	LaFayette Lakewood Lincoln
Bank Street	Tuskegee Wilmington	Tuskegee Wilmington Boulder Elmira	Tuskegee Wilmington Boulder Elmira
Oregon	Tupelo E. St. Louis	Tupelo E. St. Louis E. Las Vegas	Tupelo E. Las Vegas
Kansas	Oraibi Portageville	Oraibi Portageville Mounds	Oraibi Portageville Mounds
High Scope	Ft. Walton Central Oz.	Ft. Walton Central Oz. Greeley Seattle	Ft. Walton Central Oz. Greeley Seattle
Florida	Jacksonville Chattanooga	Jacksonville Chattanooga Jonesboro Houston	Jacksonville Chattanooga Jonesboro Houston
EDC	Washington Johnston Co.	Washington Paterson Johnston Co.	Washington Paterson Johnston Co.
Pittsburgh		Lock Haven	Lock Haven Montevideo
REC		Kansas City	Kansas City
Enablers		Billings Colorado Sp. Bellows Falls Newburgh Puerto Rico	Billings Colorado Sp. Bellows Falls Newburgh Puerto Rico

which affect that transferral -- and in the product -- the extent to which practice looks like theory at given points in time.

In the present report, we are concerned about the product of implementation as a question of internal validity. The main purpose of HSPV, as we have said, is to test the effectiveness of different educational models. Before drawing conclusions about test results, it is important to determine whether the children actually received the treatment being tested. If they did not, then we cannot assume that differences in outcomes result from the treatments and therefore cannot draw conclusions about model effectiveness. Likewise, we cannot assume that the lack of differences disprove the theory on which a model is based.

The study of the implementation process in an experimental context is important as an issue of external validity where external validity is defined as the extent to which findings can be generalized from one setting to another. One element of that generalizability is whether a treatment can be successively implemented in a variety of circumstances or whether it requires special conditions. For example, it is important for potential users to know if it takes five years or one year to implement a program or if only a particular kind of teacher works well with it.

The process of implementation is also an important topic of study in itself. We know relatively little about the process of transferring a preschool model from a sponsor's central office to a geographically distant site. It would be useful to know more in order to improve planning for future programs.

It is important to note that a major element of the study of process is knowledge of the product. While it is possible to describe the events surrounding the process of implementing a model without reference to whether the model is actually being used, and while such information may be interesting, it does not allow us to determine which factors actually affect the success of implementation and what their relative importance is.

The 1970-71 Implementation Study

This is the second report we have written on implementation in HSPV. Since that previous effort has obviously influenced this report, it seems worthwhile to summarize it briefly. Our goal in the first report was to examine both the extent and process of implementation.

Extent of implementation. The most we can say about the extent of model implementation during 1970-71, however, is that it varied from class to class within all models: some teachers did a better job of implementing the treatment than others. We cannot talk about the size of the differences -- how different the classes really are -- or about which ones are close enough to model theory to provide a valid test of the model because of the inadequacy of the available implementation measure. The measure was a simple one which asked sponsors to rate the performance of individual teachers in their model on a ten-point scale ranging from not acceptable to outstanding. The difficulty is that the findings it yields are to a great extent uninterpretable because

the standards for the rating were not made explicit. Most important, the scale was not anchored by an operational, or even agreed on, standard of full implementation. Each rater made decisions on the basis of his or her own criteria. One important lesson from the 1970-71 analysis, then, is that the measurement of implementation is not a simple, straightforward task.

Stating this lesson positively, we believe that an adequate measure of implementation must meet several criteria. The most important criterion is an operational definition of full implementation, which in turn requires two things. First, we need to describe what a fully implemented model looks like in practice. This should include both a distinction between what is and what is not essential to the model and the translation of broad goals and principles into operational terms. Second, assuming that the goal within the experimental framework is full implementation and that we need some way of distinguishing valid from non-valid examples of a treatment, we need to formulate standards for judging implementation. If model classrooms differ from one another, we need to know how much and what kinds of variation are allowable within the theory of the model and within the experimental framework of HSPV without invalidating the class's usefulness as a test of model effectiveness. As will be evident from the discussion of the 1971-72 measures, these criteria are not easy to meet.

Process of implementation. Our initial information on the implementation process came from narrative reports submitted to the Office of Child Development by consultants who had been hired to visit the HSPV sites

to monitor implementation. The consultants were experts in early childhood education and each visited the same site regularly throughout the year. The reports indicated that the implementation process in an interactive one which depends not only on the sponsor's input but on the site's reactions and characteristics. Although informative, the reports were written without reference to any common guidelines or topics and therefore could not be used for any systematic analyses.

We did have systematic data on a variety of variables taken from questionnaires completed by HSPV teachers and aides, sponsors and OCD consultants and we attempted to explore the implementation process by using these variables in regression analyses. Although these data were systematic, they are far from satisfactory because many of the questionnaire items were ambiguously worded, and more important, there were no items for a number of variables we would have liked to test. For example, we suspected that the Head Start directors played an important role in HSPV; their understanding of and attitudes toward the model probably influenced the staff's implementation. Yet we had very little data on the directors and could not test that relationship. The dependent measure was also problematic since the only measure available was the sponsor ratings, whose limitations have already been described. Obviously the resulting analyses could be described, at best, as exploratory. Given all these limitations, we were surprised to find an equation which explained almost 98% of the site-to-site variaiton in sponsor ratings. The equation contained

only three independent variables: the amount of intra-staff friction in the site, the rapport between the administration and the staff, and the adequacy of the physical plant of the site -- all as judged by the sponsors. The interpretation of this finding is not clear: we cannot tell whether pleasantness in a site is a factor which actually influences implementation or whether the sponsor simply equates pleasantness with high implementation. We will return to this finding when describing this year's report. Our attempts to account for within-site variance in sponsor's ratings were less successful but did support the general premise that model implementation was affected by factors in the site as well as by the input made by the sponsor.

The main conclusion of the 1970-71 report was that the process -- and the study -- of implementation were more complex than we anticipated and that we had not succeeded in dealing adequately with the questions we had raised about the process and extent of implementation.

The 1971-72 Analysis:

Our initial goal for this analysis, then, was to ask the same questions that we did in last year's report, but with hopefully better data. One advantage of the third year data over the second year data is that Huron designed four new measures of the extent of implementation. Although these new measures are better than the sponsor ratings, they also have problems. Consequently, we can conclude that there is variation among classrooms

in levels of implementation but again, we cannot make any statements about the size of the differences among classrooms or about their relation to full implementation. Also, a preliminary attempt to relate some of the implementation measures to outcome did not reveal any systematic patterns.

In an attempt to obtain better information on the implementation process, Huron revised the existing questionnaries by clarifying, adding, and eliminating items from the teacher and aide questionnaires. We also designed a new questionnaire for the Head Start directors. We found, however, that the data were still inadequate for fully explaining and understanding the implementation process. Every question that was answered in the questionnaire seemed to raise several others that we had not asked. Further, after doing numerous regression analyses, we found that the equations changed completely when different measures of implementation were used as the dependent variables. It seemed unwise, to say the least, to draw conclusions about factors which influence implementation when these relationships are almost completely dependent on the measure they are used with.

One equation which should be mentioned specifically, however, is the replication of the equation from the second year analysis which demonstrated the almost perfect relationship between the site average ratings of implementation and the pleasantness of a site. Because we did not use the form from which the sponsor's ratings of site variables other than implementation was taken in the third year, the analysis was not replicated exactly. One would expect, however, that if there is a real relationship

between these site characteristics that it would show up at least moderately when judgments by other people on similar variables were made. The new equation, however, showed almost no relationship. It appears that the inital equation was a function of sponsors' ~~not distinguishing~~ failure to distinguish between pleasantness and implementation, though the direction of causality is not clear.*

With this exception, the replication and updating of the 1970-71 analyses did not change the main conclusions of that report and did not shed any new light on the product or process of implementation.** Rather than present the replication in detail, then, we decided to concentrate on one aspect of the study of implementation for this 1971-72 report: the development of measures of implementation. In this way, we hope to document what we have learned in the process of developing, using, and analyzing data from instruments to measure the extent of implementation.

*A more detailed presentation of this analysis can be found in Appendix A.

**All the analyses for the replication actually were completed. For example, analyses of variance were done on every implementation measure to test for the effects of models and sites within models and, in addition, for the sponsor ratings to test for the effects of time of ratings. Basic statistics were calculated and analyses of variance testing model and site effect were performed on a variety of variables from the teacher questionnaire and the directors' report which were thought to be important in predicting levels of implementation; correlations among these measures were also calculated, a number of regression analyses were done to account for variation in the extent of implementation. Items from the teacher and director questionnaire were used as independent variables and the sponsor ratings and the consultant checklist averages were used as dependent measures in separate equations. One series of regressions were undertaken to test hypotheses about the relative importance of factors which influence implementation. On finding that the resulting equations differed considerably with the different dependent measures, another series of analyses was done in attempt to explain the differences. Still another series was performed to explore differences in equations for structured and unstructured models.

In the next section we will examine HSPV implementation measures individually: their design and use will be described, their strengths and weaknesses will be discussed, and the findings obtained from them will be presented briefly. In the following section we will consider the relationships among the measures: because the measures still tend to be subjective, we are interested in learning whether there is any agreement among different raters in their judgments about the extent of implementation. In the final section of the report, we will summarize our conclusions about measuring the extent of implementation and about HSPV in general.

09013

MEASURES OF IMPLEMENTATION

In the 1971-72 study, there were five measures of implementation. The major new measure was based on model-specific checklists which were designed for use by the OCD consultants. A second measure was developed by condensing these consultant checklists to a shortened form to be used by the sponsors. These sponsor checklists were accompanied by a request for an overall rating of each classroom by the sponsor using the same response scale as used in the checklists. These overall ratings provided a third measure of implementation. A fourth measure was taken from a questionnaire completed by the HSPV teachers.* The fifth measure was the same sponsor rating of teachers that was used in last year's analysis. The means and standard deviations for sites and models in each of these measures are shown in Table 2.**

Consultant Checklists:

The consultant checklists were designed to systematize the OCD consultants' observations of model implementation. As we explained earlier, the OCD consultants were people who were hired

*An item similar to this one was completed by the Head Start directors but is not included as an implementation measure because there was only one rating for each site and the variation in responses was slight.

**Copies of the measures can be found in Appendix B.

Table 2. Means and standard deviations by site and by model on five measures of implementation completed in the spring of 1972.

Model	Site	Consultant Checklist			Sponsor Checklist			Sponsor Overall Rating			Teacher Perception			Sponsor Rating		
		mean	s.d.	N	mean	s.d.	N	mean	s.d.	N	mean	s.d.	N	mean	s.d.	N
Far West	Buffalo	2.94	0.69	5	2.94	1.13	5	2.00	0.71	5	2.09	0.70	11			
	Duluth				1.98	0.61	5	3.00	0.71	5	2.40	0.55	5	3.80	2.39	5
	Salt Lake	2.03	0.67	3	2.88	0.68	5	3.00	0.0	6	2.60	0.55	5	5.20	1.79	5
	Tacoma	3.72	0.19	6	3.48	0.27	6				2.83	0.75	6	6.86	0.90	7
	Overall	3.08	0.82	14	2.89	0.89	21	2.69	0.70	16	2.41	0.69	27	5.39	2.03	17
Arizona	Lafayette	2.50	1.37	5				2.58	0.38	6	2.67	0.82	6	6.33	1.51	6
	Lakewood	3.90	0.0	1	2.47	0.56	6	2.67	0.52	6	2.50	0.58	4	6.00	1.69	8
	Lincoln	2.70	0.46	3	2.03	0.27	6	2.33	0.82	6	2.67	0.52	6	4.50	1.64	6
	Overall	2.72	1.09	9	2.25	0.48	12	2.53	0.58	18	2.63	0.62	16	5.65	1.73	20
Bank Street	Boulder	3.95	0.13	4	2.90	0.42	4	2.88	0.25	4	2.25	0.50	4			
	Tuskegee	3.55	0.24	4	3.06	0.45	5	2.70	0.27	5	2.75	0.62	12	5.40	1.14	5
	Wilmington	2.87	0.25	6	2.25	0.68	6	2.50	0.78	6	2.67	0.49	12	5.17	0.75	6
	Elmira	2.83	0.10	4	2.56	0.61	5	2.60	0.55	5	2.80	1.10	5	5.80	1.48	5
	Overall	3.25	0.51	18	2.71	0.58	20	2.70	0.50	20	2.67	0.65	33	5.44	1.09	16
Oregon	Tupelo				2.20	0.26	4	4.00	0.0	3	2.75	0.50	4	5.50	2.08	4
	E. Las Vegas				2.66	0.52	5	4.00	0.0	5	3.60	0.89	5	6.00	1.87	5
	Overall				2.46	0.47	9	4.00	0.0	8	3.22	0.83	9	5.78	1.86	9
Kansas	Oraibi	3.38	0.26	5							3.29	0.95	7			
	Portageville	3.30	0.0	4							2.25	0.96	4			
	Mounds										2.20	0.45	5			
	Overall	3.34	0.19	9							2.69	0.9[illegible]	16			
High/Scope	Ft. Walton Beach	2.90	0.17	5	2.64	0.13	5	3.00	0.0	3	3.00	0.0	5	6.60	1.34	5
	Central Ozarks	2.00	0.45	5	2.38	0.13	5	2.60	0.55	5	2.81	0.66	16	6.00	0.89	6
	Greeley	2.95	0.07	2	2.65	0.21	2	2.75	0.35	2	2.50	0.58	4	5.75	0.96	4
	Seattle	3.16	0.38	5	2.42	0.15	6	2.50	0.55	6	2.50	0.84	6			
	Overall	2.72	0.58	17	2.49	0.18	18	2.69	0.46	18	2.74	0.63	31	6.13	1.06	15
Florida	Jacksonville				2.74	0.45	4	2.75	0.50	4						
	Jonesboro	4.06	0.48	3	2.69	0.48	4	2.75	0.50	4						
	Chattanooga	3.90	0.38	2	2.50	0.51	12	2.50	0.52	12						
	Houston	3.64	1.03	3	2.88	0.34	9	2.89	0.33	9						
	Overall	3.75	0.53	8	2.68	0.44	29	2.69	0.47	29						
EDC	Washington										1.50	0.71	2			
	Paterson	3.17	0.29	3							2.00	0.0	4			
	Johnston Co.	2.55	0.38	4							3.20	0.45	5			
	Overall	2.81	0.46	7							2.46	0.82	11			
Pittsburgh	Lock Haven	3.00	0.16	4	3.20	0.59	6	3.00	0.63	6	2.71	0.49	7	5.33	0.82	6
	Montevideo	3.20	0.0	1	2.80	0.0	1	3.00	0.0	1	2.50	0.71	2	5.50	0.58	4
	Overall	3.04	0.17	5	3.14	0.56	7	3.00	0.58	7	2.67	0.50	9	5.40	0.70	10
RLC	Kansas City				2.76	0.35	6	2.00	0.0	6	2.25	0.50	4	6.67	0.52	6
Enablers	Billings										2.43	0.54	7	5.50	0.87	6
	Colorado Springs										3.40	0.55	5	6.50	1.87	6
	Bellows Falls										2.75	0.50	4	5.75	2.22	4
	Newburgh										3.00	0.0	4	6.00	1.55	6
	Puerto Rico										2.33	0.52	6	7.00	0.89	6
	Overall										2.73	0.60	26	6.18	1.49	28

Note: The numbers in this table are carried to two decimal places in order to demonstrate that there is some variation in the way sites are rated; if they were rounded, as they should be given the values of the raw data, these differences would be eliminated.

The data were gathered with measures with which are model specific, either by design, or by default because no rating standards were made explicit. Therefore, they should not be used to compare levels of implementation in different models.

9016

by the Office of Child Development to monitor implementation in HSPV sites. The reports written by these consultants during the first two years of the study, although rich in anecdotes, did not provide information appropriate for systematic analysis. In an attempt to systematize at least part of their reporting, Anne Monaghan of the Huron Institute, in conjunction with the consultants and sponsors, developed checklists to guide the consultants' observations of model implementation. Each checklist contains a description of the model, in the form of a list of component items. The consultants, after observing a classroom, rated each component item on a scale of 1 to 4 (1 - not at all implemented, 2 = implemented to some extent, 3 = implemented to a great extent, 4 = technically implemented).* The design, administration, and patterns of ratings on individual components from the checklists can be found in a separate report by Monaghan.** Here we will deal with the measures only in their summary form.

The strategy of the checklist approach was to describe each model in terms of its individual components and to rate the implementation of the components rather than the model as a whole. Since most of the available literature contained only partial or general descriptions of the models, the sponsors themselves were

---

*Some items had a scale of 1 to 5 where 5=beyond technical implementation; this category was intended for staff members who not only implemented an item but were expanding on it or using it creatively.

**Monaghan, Anne, An Exploratory Study of the Match Between Classroom Practice and Educational Theory: Models in Head Start Planned Variation. Doctoral dissertation, Harvard Graduate School of Education, 1973.

asked to articulate, as clearly as possible, what a fully implemented classroom should look like. Did the model specify particular materials, organization of space, daily routine and schedule? Did the model specify parent or paraprofessional roles in the classroom? Checklists were to include only those specifications of the models which were considered essential for full implementation. Thus, the checklists were attempts to describe an ideal model classroom in operational terms.

This checklist approach is an improvement over the sponsor ratings used in the previous report, first, because it anchors the scale in the sense of making the basis for the judgments more explict: a small unit of behavior or single principle is easier to judge clearly and to agree on than is an entire model. Second, the checklist approach specifies which components are implemented and which are not. Such information is useful, for example, in determining whether the model as implemented in two classes is similar enough for them to be considered to be examples of the same treatment.

The actual process of developing the checklists, however, revealed problems in using the approach in HSPV. As we attempted to examine each model's theory in order to create operational definitions of each treatment, we realized that many model theories were not easily translated into operational specifications for classroom practice. Most of the HSPV models did not have detailed specific goals, such as teaching children colors, shapes, numbers,

the alphabet, nor prescribed methods for reaching their goals. In fact, some models were theoretically opposed to translating their goals and principles into closely defined behavioral specifications. Rather, the implementation of those models required creative adaptation of philosophical principles of educating children. An example of such a principle is that children should be encouraged to make their own discoveries, set their own pace of learning, choose their own activities, The teacher's role in this situation is to use his or her own intuitions and good judgment about how to encourage the children. The sponsor's role is to develop the teacher's own abilities in teaching young children rather than to instruct the teacher as to how to teach specific lessons.

As a result of these characteristics, some model components, particularly those in models with developmental rather than behaviorist theories, could not be stated in terms of specific, easily observable behaviors. For example, a Bank Street item read: "The adults provide opportunities for good relationships between child and adults as well as between children and children." In order to rate this item, the observer had to know what was meant by "good relationships" and by "provide opportunities." This item could be interpreted in various ways by different observers depending on their values. Further, it was not explicit how many and what kinds of "opportunities" had to be provided in a classromm observation period in order for the observer to rate the item "4"

(technically implemented). The problem of subjectivity in rating implementation, then, was not completely eliminated by the checklists.

Another implication of models operating on general principles rather than carefully specified behavior is that classes -- even if they were fully implemented -- would not be identical replications. Indeed, if teachers were to be creative and spontaneous, they could hardly have been required to conform to specific behavior patterns and routines. Moreover, where models encouraged teachers to create opportunities for each individual child to learn what was relevant to him or her, each child in the classroom in a sense received a different treatment. Therefore, the treatment could vary at the child level, the classroom level, and the site level. Replication, then, is an inappropriate standard for determining whether a class represents a valid example of the treatment to be tested. The design of the checklists, however, did not meet the problem of establishing an alternative standard. There are no means for deciding how much variation is allowable within a single treatment nor how many components have to be present for a model to be considered implemented.

Thus, the checklists do not meet the experimental requirements of operationalized treatments and standards of implementation. While this discrepancy arose to a large extent because of the nature of a number of the models, it should be emphasized that models based on general principles are not necessarily harder to implement

or less theoretically sound than models which specify behavior. They are simply less amenable to study within a large-scale experiment based on a classical design.

Another set of problems with the checklists stems from the fact that there was neither time nor money to field test the checklist instruments before their use in 1971-72. A small scale reliability study was done in mid-year and showed mixed findings; some checklists and some items appeared to be more reliable than others, but different methods of computing measures of reliability yielded different results. Moreover, not all of the consultants received systematic training, either in model theory or in observational methods and the use of the instrument. Most of the information about the use of the checklists was conveyed by telephone or by letter. Not only was there no assurance that consultants had common assumptions or standards in using the instruments, but it appears that there was some variation in the extent to which the consultants even understood the models and the instruments.

Hence, the consultant checklists, are an improvement over the sponsor ratings but they still have severe limitations. The ratings often require subjective judgments and the response scales are ordinal at best since we did not know that the intervals between ratings are equal. Because of the lack of training of the consultants and because each site was rated by a different person, it is doubtful that the data should be used to compare across sites within the same model. Moreover, because the check-

lists are model specific, they should not be used to compare the extent of implementation across models. The primary value of the checklists, then, seems to lie not in the data which were collected, but in the attempt to translate educational theories into operational descriptions of classroom practice.

Data analysis. The checklist data reported here are simple means of all items rated on the checklists completed between April and June of 1972. We include only one checklist average per classroom. Although the time restriction eliminated some classrooms, we decided that a spring 1972 measure would be more comparable to the other four measures of classroom implementation which were also completed in the spring of 1972.

There are three models for whom we have no checklist data. The Enabler model consisted of five early education specialists, each working in a separate community to help the community develop its own model. Because the model was not created until the second year of HSPV and because the basic premise of the approach is local program development, the Enabler consultants had not reached a consensus about the common components of the approach when the checklists were being developed. Therefore, no measure exists for this model. In the Oregon model, the checklists rated in the spring were not complete. At one Oregon site, the classrooms were preparing for a special event and thus were not using the model curriculum. The REC model is not included because there were checklist data for only two teachers.

The column of means and standard deviations for the consultant checklist in Table 2 shows that there is a good deal of variation among sites within the same model on this measure. The most striking example is the Far West model where the Tacoma mean is 3.72 (a nearly perfect score) and the Salt Lake mean is 2.03. We suspect that this variation is at least in part a function of hhe fact that there were different observers for each site. Observers may have used the response categories differently; some observers, for example, may have been more reluctant than others to rate items as fully implemented. The variation among classrooms within the same site tends, with the exception of LaFayette in the Arizona model where the standard deviation is 1.38, to be smaller than the site-to-site variation. The small standard deviations do not necessarily indicate that there is no variation within sites: since the response categories are limited, there is not much room for variation. Moreover, because all classes within a site are rated by the same person, we can conclude that the variation which exists indicates that the raters are distinguishing among classes in the extent to which they are implemented.

Sponsor Checklists:

The sponsor checklists were condensed versions of the consultant checklists, and were also rated on a 1-4 scale (not at all implemented to technically implemented). Like the consultant checklists, there were different instruments for each model, and

the data from them cannot be used to compare models on levels of implementation. Also, as derivatives of the consultant checklists, they shared the problems of operationally defining models and setting standards for assessing the validity of the treatments. The advantage of the measure is that the ratings were done by people who supposedly know the model best. If this is true, then the problem of items not being operationally defined may not be as serious on the sponsor checklist as on the consultant checklist because the sponsors should know what their model should look like in practice. Sponsor staff also had the advantage of perspective over more than one site. Thus, there should be more consistency among sites in the way the ratings are used, and as a result, the variation among sites within the same model should not be as much a result of different rater as on the consultant checklists.

There are several problems, however, which affect the sponsor checklist more than the consultant checklist. One is the possibility of sponsor bias. Because sponsors had an investment in the success of the implementation effort, they may either have overrated the total staff's performance or they may have been overly critical. A second problem is that the sponsor checklists were not rated immediately after observation of the classroom; the ratings were based more on memory of past visits to the site. Thus, their ratings may have been less accurate than the consultant ratings.

Another problem resulted not from the nature of the instrument, but from a misunderstanding about who was to complete it. The instructions stated that sponsors should "ask the person on your staff who is most knowledgeable about a particular site to complete the forms for that site". The instrument was designed to be rated by staff in the sponsor's central office, but unfortunately, some sponsors forwarded the instruments to the sites, where local site staff completed them. These sponsors undoubtedly felt that the people most knowledgeable about a site were people at the site. However, when the sponsor checklists were completed by site staff, they lost some of their intended value as judgments of sponsors. When local site staff rated the checklists, the problems of the instrument became more similar to the problems connected with the consultant checklists.

Data analysis. The sponsor checklists were completed in May, 1972. There are data for eight models. The EDC and Kansas sponsors did not complete the forms because they felt that rating classrooms and teachers was inconsistent with the model theories. The rating of classrooms and teachers was believed to undermine the supportive relationship that should exist between sponsor and teacher. Enablers were not included for the same reason they were not included in the consultant checklist analysis: no checklist was ever developed. In our analysis, we use simple means of all the teacher behavior item ratings on the sponsor checklist.

The column of means for the sponsor checklist in Table 2 shows fairly low scores. Only three sites (Tacoma in the Far West model, Tuskegee in the Bank Street model, and Lock Haven in the Pittsburgh model) received ratings over 3.0. In all but two cases, the sponsor checklist rating is lower than the consultant rating. It appears that sponsors did not bias their judgments to enhance the appearance of high levels of implementation, but in fact, may have applied stricter standards than the consultants. The variation among sites within the same model is not as great as on the consultant checklist, supporting our contention that some of the site variation on the consultant checklist may have been a result of the fact that each site had a different observer. Finally, the sponsors also tend to distinguish among classrooms within sites in making their ratings.

Sponsor Overall Ratings:

The sponsor overall rating is a global judgment of the level of implementation in the classroom, rated on the 1 to 4 scale. The request for these ratings was included in a covering letter sent with the sponsor checklists in May of 1972. Sponsors were asked to rate the classroom on overall implementation after completing the checklist for that classroom.

The sponsor overall rating was designed to provide a "round number" evaluation which was anchored by the component items on the sponsor checklist. It was reasoned, for example, that a classroom could hardly receive a "4" (technically implemented) overall rating unless a large proportion of the component items on the

checklist were rated "4". As sponsor judgments, the overall ratings have the same advantages as the sponsor checklists: perspective over more than one site and knowledge of the model.

As we mentioned in the previous section, there was some confusion about who was to complete the checklist and overall ratings. Some sponsors did not complete the measures, but forwarded them to the site where local personnel completed them. In a few cases, sponsors separated the two measures, sending the checklist to the site, but making the overall ratings themselves.* Therefore, the attempt to obtain an anchored overall sponsor judgment was not entirely successful.

Data analysis. There are sponsor overall ratings for the same eight models for whom we have sponsor checklist data. The means on the sponsor overall rating are generally low. An exception is the Oregon model which rated all of its classrooms as technically implemented (4.0). The meaning of this finding is unclear, however, since the scores on the sponsor checklist for this model ranged between 2.20 and 2.66. All of the site means in the other models on the sponsor overall rating range between 2.0 and 3.0. There is very little variation among sites within the same model on this measure: the range among sites is no more than 0.5 except in the Far West model where Duluth scores a full point lower than Salt Lake and Tacoma. Perhaps the global nature of the

*In a few cases, the sponsor checklist average is different from the sponsor overall rating by as much as 1.5 on a 4 point scale

sponsor overall judgment and the limited number of response categories obscured differences among sites.

Teacher perception. Teacher perception ratings were taken from an item on a questionnaire distributed to all HSPV teachers in the late spring of 1972. The item read:

> How different do you think your classroom is from what the sponsor says an ideal classroom for your model would be? (Check one category...)
>
> 1 = very different
> 2 = somewhat different
> 3 = not at all different
> 4 = don't know

While teachers were probably the best source of information about actual classromm activity, they did not have the advantage of observing implementation in other sites, or even necessarily in the other classes in their site, and do not have the advantage of comparisons to give better perspective over the model. They may not have had as much knowledge of model theory as the sponsor, or in some cases as much as the consultant. They could hardly have been objective, since, in one sense, they were asked to judge themselves as implementers.

A major problem with this instrument is the number of possible interpretations of the teachers' judgments. If a teacher rated his or her classroom as "somewhat different", we do not know from this measure alone whether that response indicates partial implementation as it was intended to or whether it indicates that the teacher fully implemented and then moved on to extend it creatively

as might be suggested by the fact that some of the teachers who chose this response were rated high on the sponsor rating to be described in the next section. Similarly, if a teacher judged her or his classroom to be "not at all different" it might indicate full implementation as was intended. It is also possible that it might indicate lack of understanding of the model's standards or over-confidence in cases where the teacher is rated comparatively low on other measures.

It is worth noting that 86% of the teachers completing this item checked either 2 (somewhat different) or 3 (not at all different). Although this measure was not designed to yield a dichotomous response, it does, in fact, separate teachers into two groups, based on a fairly minor distinction between "somewhat different" and "not at all different".

Data analysis. There are teacher perception ratings available for 221 out of the total 247 tested HSPV classrooms. Thus, the problem of incomplete data, which applies to the other measures, is not a serious difficulty with the teacher perceptions. The Florida model was excluded from these analyses because the teacher in that model is not the primary agent for innovation; the parent educator is.

The means and standard deviations for the teacher perception ratings in Table 2 reveal a number of interesting findings. One is that all of the teachers in Ft. Walton Beach (in the High/Scope model) stated that their classrooms were not at all different

from the ideal model. Thus, these teachers were confident that they were implementing the model. However, the other measures for this site are fairly low. Perhaps these classrooms were underrated, or perhaps the model theory was not thoroughly understood by these teachers. Another interesting finding is in the EDC model: all of the four teachers in Paterson checked "somewhat different" while all but one of the five teachers in Johnston County checked "not at all different" (one checked "don't know"). Such nearly perfect site variation rarely occurs and may reflect a systematically different in sponsor input to the sites. A third finding is that four out of five teachers in East Las Vegas (Oregon model) stated that they didn't know how different their classrooms were from the ideal model (one teacher checked "somewhat different"). Yet, they were all given the highest score possible on the sponsor overall rating. Perhaps the teachers were simply too modest to rate themselves high, or perhaps the sponsor had not given them enough feedback to know that they were doing well.

Sponsor ratings:

The sponsor rating form was sent to each sponsor twice during the 1971-72 school year: once in November and again in May. This is essentially the same form that was used as the primary implementation measure during the first two years of HSPV. It requested an evaluation of each teacher as to her or his performance in the model on a scale from 0 to 9 (0=unacceptable; 1-3=barely acceptable; 4-6= average; 7-9= outstanding).

There are two main problems with this instrument. First, the response scale implies a continuum but it has no anchor in a definition of full implementation. The judgment of "outstanding" implies that a teacher must be better in some way than other teachers, but it does not tell us whether that teacher is fully implementing the model. It may be that the outstanding teacher's classroom was not fully implemented, but only came closer to full implementation than the other classrooms. Similarly, the term "average" implies that a teacher is in the middle -- better than some teachers, not as good as others. It is not clear whether an average performance is a good one (if most teachers are good at implementing the model) or a fair one (if most teachers were only fair). Such a response scale suggests that sponsors were "graded on the curve" in the sense of measuring one teacher against others rather than against a fixed standard.

Second, there was no information about how or in what areas a teacher was deficient or outstanding: an "average" teacher may have been outstanding in some areas and unacceptable in others, or she may simply have been average in all areas.

As a result of these problems, the sponsor ratings data can be used to make statements about a teacher's performance in relation to other teachers, but they cannot be used to indicate either how close that performance is to full implementation, which components of a model were implemented, or whether one model was better implemented than another.

The main value of this instrument is that ratings were made by the people who knew the model theory and standards best: the sponsors and their field representatives. Another advantage of these data over the other implementation measures from the researcher's point to view is that there are data for four points in time spread over two years: the forms were completed twice during the year 1970-71 as well as twice during 1971-72. With a sub-sample of teachers who were rated at all four points, we can examine patterns of implementation over time.

Data analysis. There are sponsor ratings data for eight of the eleven models. As with the sponsor checklist and overall rating, EDC and Kansas did not complete the forms. The Florida model was excluded for the same reason it was excluded from the teacher perception analysis: the measure focuses on the teacher, but the teacher does not have primary responsibility for implementation in this model.

The data for the sponsor ratings for spring 1972 shown in Table 2 indicate that the means for all models range between 5.0 and 7.0 (middle average to low outstanding). There is some variation among sites within models, especially in the Far West model. The variation among sites seems to be fairly consistent with three of the other measures (the teacher perception does not seem to be consistent). In most cases, where a site is rated quite low (or high) on one measure, it is also rated low (or high) on the other measures.

In order to examine patterns of implementation over time we selected a subsample of teachers for whom there were data on the sponsor ratings at all four points: February and May of 1971 (the second year of HSPV) and October of 1971 and May of 1972 (the third year of HSPV). Although this decision severly limits the sample size, it was necessary in order to eliminate the explanation that trends -- or the lack of them -- resulted from teacher turnover.

The hypothesis we wanted to test with these data is that implementation improves consistently over time. It is commonly assumed that given additional training and exposure and time to work with a model, a person will use the model to a greater extent than before. The data from the sponsor ratings do not support this hypothesis. The model means in Table 3 and the graph in Figure 1 show no consistent trend over time. Most model means tend to go up in May 1972, but since that change is not part of a longer trend, it seems probable that the final point is more indicative of an upward bias in the sponsor's judgments than a sudden burst in the teachers' performance. The same lack of clear patterns tends to hold for individual teachers.* Although some teachers, particularly those who were quite low on the first rating, improved from February 1970 to May 1972, the scores for most teachers simply moved up or down from rating to ratings by one or two points.

---

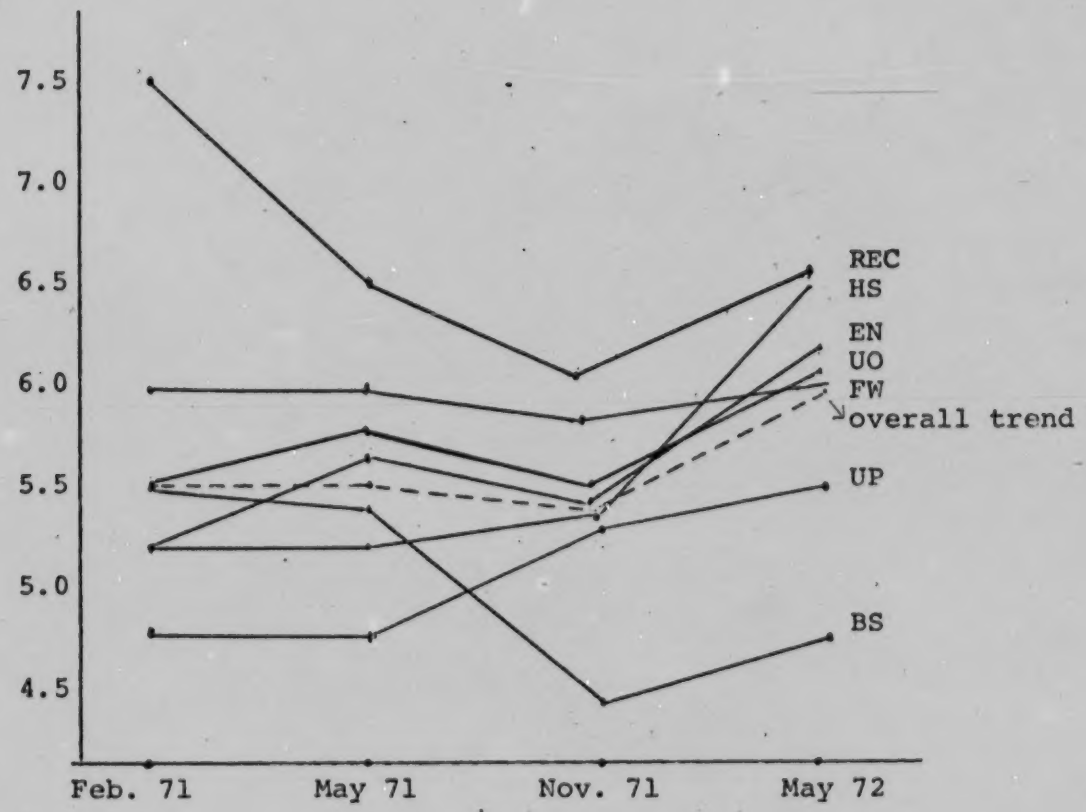
*The sample is too small to report site means.

Table 3. Model means on sponsor ratings for four points in time.

Model	N	Feb '71	May '71	Nov '71	May '71
Far West	8	5.9	5.9	5.8	5.9
Bank St.	8	5.5	5.4	4.5	5.0
Oregon	6	5.5	5.7	5.5	6.0
High/Scope	6	5.2	5.2	5.3	6.5
Pittsburgh	5	4.8	4.8	5.2	5.4
REC	2	7.5	6.5	6.0	6.5
Enablers	11	5.2	5.6	5.4	6.1
Overall	46	5.5	5.5	5.3	5.9

00034

Figure 1. Model means on sponsor ratings for four points in time.



Although these analyses throw doubt on the hypothesis of improvement over time, they do not completely disprove it. While it may be that the lack of clear trends over time indicates that factors other than exposure to the model influence the extent to which it is implemented, the finding may also be an artifact resulting from measurement problems. Because the rating scale is not anchored in explicit criteria, a sponsor's use of the scale may change from time to time. In such a case a decrease in reported level of implementation may not correspond to a real change in behavior but only to the unreliability of the instrument.

Summary:

On all implementation measures we found some variation among classrooms within models. Although limited by small ranges on four of the five response scales, it does appear that raters tend to distinguish among staff members in the extent to which they implement the models. The data also show that the means for all models on all measures tend toward the middle of the scale. While it is encouraging to find that on the average, people seemed to do fairly well at implementing the programs, we do not find uniform excellence.

As with last year's findings, our conclusions cannot go much beyond these statements. We do not know how large the class-to class variation is in terms of actual classroom operation, nor do we know which classrooms are close enough to the model ideals

to serve as valid tests of the model theory. Moreover, although Monaghan's work took an important step toward specifying the models through the consultant checklists, most of the models are not described in operational terms, and no definition of full implementation has been developed to replace the standard of identical replication.

10037

CORRESPONDENCE AMONG IMPLEMENTATION MEASURES

Given the problems of the existing implementation measures -- the lack of anchored scales and the subjectivity of the ratings -- and given the argument that different raters judge the extent of implementation from different perspectives, it seems reasonable to ask whether there is any correspondence among the different measures of implementation. Are ratings similar from measure to measure, or is each one different? Would we draw the same conclusions about the extent of implementation from all of them?

If there is little agreement among me[illegible] -- if, for example, a teacher is rated high by one pe[illegible]nd low by another -- we would conclude that implementation ratings are idiosyncratic and that the measures are of little value in systematically assessing this phenomenon. In fact, if different raters have very different judgments about who is doing a good job of implementing a model, then we would question whether the label, implementation, has any shared meaning. In contrast, if there is substantial agreement among different raters (at least in their ranking of people if not in the assignment of particular numbers), then we would have confidence that the implementation measures despite their problems have some value.

Since scanning the site means in Table 2 does not give us an adequate answer to the question of correspondence among

measures -- both because the table shows both consistency and inconsistency and because it gives us no information about ratings within sites where we would expect the agreement to be highest -- we have done two additional analyses.

The simplest way of looking at the correspondence among measures is to examine the product - moment correlations between any two of them. The sample size for the correlations varies from about 60 to 90, depending on the pair of ratings. This is smaller than the sample sizes for the analyses of the measures separately because it requires a match between two points; if data are missing from either points, the case must be discarded. We do not report significance levels because it is mathematically impossible to attain perfect correlations with these data because they have neither normal distributions nor homogeneous variances. The correlations, then, should be used only as heuristic indications of the relative agreement between the various implementation ratings.

Table 4 shows the correlations between ratings paired at the classroom level. The three highest correlations are for the three pairs of sponsor judgments, as might have been predicted: the sponsor ratings correlate .61 with the sponsor checklists and .48 with the sponsor overall ratings, and the latter two correlate .53 with each other. Because these three instruments were completed, in most cases, by sponsor staff representatives, they should be highly correlated, higher perhaps than they

Table 4. Product-moment correlations between pairs of ratings.

	Teacher perception	Sponsor rating	Consultant checklist	Sponsor checklist	Sponsor overall rating
Teacher perception	1.00	-.04	-.05	-.02	.14
Sponsor rating		1.00	.41	.61	.48
Consultant Checklist			1.00	.46	.23
Sponsor Checklist				1.00	.53
Sponsor overall rating					1.00

Note: The N's for the correlations range between 60 and 90.

20649

actually are. The next highest correlations are between the consultant checklist ratings and sponsor judgments (.41 with sponsor ratings, .46 with sponsor checklists, and .23 with sponsor overall rating), indicating that the consultants and sponsor staff agree with one another to some extent. The teacher perception ratings do not correlate highly with any of the other measures. This finding is not surprising: the teacher perception instrument was quite different in design from the other instruments. In addition to the fact mentioned earlier that the teachers were in the unique position of rating themselves, they also were unique in only rating one classroom. On the sponsor and consultant measures, the same person rated all of the classrooms at a site, and thus the rater was in a position to discriminate among classrooms. Had teachers been asked to rate all of the classrooms at their site, it might have been more likely that their ratings would correlate with the other measures. It appears, then, that the sponsors and consultants to some extent have a common idea of how teachers should be ranked in terms of the extent to which they implement the models.

Another approach to examining the correspondence between implementation is to look at the relationship between all five measures together rather than between only pairs of measures.

19041

Since we are interested in determining whether different imstruments are measuring the same thing, it is logical to perform a factor analysis. The purpose of a factor analysis is to reduce a large number of operational variables, in this case implementation ratings, to a smaller number of conceptual variables. The assumption is that if ratings are correlated, then, the relationship may be due to one or more underlying factors which accounts for a large proportion of the variance. In the present case, finding one common factor would indicate that raters have some shared conception of what it means to implement a model. Finding several factors would suggest that they do not.

The first step in the factor analysis is to inter-correlate the five variables and subject them to a principal components analysis. The purpose of a principal components analysis is to determine the set of loadings which maximizes the correlation between the original variables and the factors.

The second step in the factor analysis is to rotate the components according to the Varimax criterion. The reason for rotation is to simplify interpretation by finding a set of variables that are highly correlated with some factors but uncorrelated with others. A factor loading can be seen as the relationship between one variable and one factor. The cluster of variables which loads high on a factor are used to describe it.

The factor analysis was performed on only 35 classrooms. The problem of incomplete data, mentioned in the previous sections, is magnified to such an extent when we attempt to look at all five measures together that there are complete data -- ratings on all

five measures -- for only these few classrooms (from a total of 257). The 35 include classrooms from only 5 of 12 HSPV models, and from only 10 of 28 sites. Since the subsample was not randomly selected, it cannot be considered as a representative sample of all HSPV classes. It does not seem to be biased, however, in terms of either systematically higher ratings, ranges of scores, or length of teacher experience with the model. Table 5 displays the data for the subsample.

Table 6 presents the rotated factor loadings on two factors. The first factor shows high positive loadings for four of the five variables. Only the teacher perception stands out as unrelated to the factor. The three sponsor judgments are very high at .86 or better, and the consultant rating is still quite respectable at .67. On the second factor, the teacher perception is the only high loading. The interpretation of these loadings is similar to the findings of the correlations: four of the five measures are highly related to one another. The teacher perception rating is, however, orthogonal to the other four variables.

Another logical question which is not dealt with by these analyses is whether there is more agreement among implementation measures in some models than in others. Although the data in Table 2 suggest that there might be differences among models in agreement at the level of site averages, formal analyses of those differences or differences (agreement) at the classroom level were not possible because of the small number of cases within each model.

In summary, the hypothesis that the various instruments tap a common notion of implementation is confirmed in part by both

Table 5. Scores on five measures of implementation for subsample of 35 classrooms with complete data.

Model	Site	Class	Consultant Checklist	Sponsor Checklist	Sponsor Overall Rating	Teacher Perception	Sponsor Rating
Far West	Salt Lake	A	1.3	2.0	2.0	3	4
		B	2.2	3.6	4.0	2	8
		C	2.6	2.8	3.0	3	6
	Tacoma	A	3.8	3.5	3.0	3	6
		B	3.7	3.3	3.0	3	8
		C	3.8	3.3	3.0	3	6
		D	4.0	4.0	3.0	2	7
		E	3.5	3.3	3.0	2	8
Arizona	Lincoln	A	2.3	1.7	1.0	2	2
		B	2.6	2.1	3.0	3	5
		C	3.2	2.3	3.0	3	6
		D					
Bank St.	Tuskegee	A	3.3	3.0	2.5	2	5
	Wilmington	A	2.7	2.4	2.5	3	5
		B	2.8	2.0	2.5	2	5
		C	2.8	2.7	3.0	3	5
		D	3.0	2.5	3.0	3	6
		E	2.6	1.0	1.0	3	4
		F	3.3	2.9	3.0	3	6
	Elmira	A	2.9	2.7	3.0	2	6
		B	2.9	2.1	2.0	2	5
High/ Scope	Ft. Walton Beach	A	3.0	2.7	3.0	3	8
		B	3.0	2.7	3.0	3	8
		C	2.9	2.7	3.0	3	5
		D	3.0	2.4	3.0	3	6
		E	2.6	2.7	3.0	3	6
	Central Ozarks	A	1.7	2.4	2.0	3	5
		B	1.6	2.3	2.0	3	5
		C	2.2	2.6	3.0	3	6
	Greeley	A	2.9	2.8	3.0	2	6
		B	3.0	2.5	2.5	2	5
Pittsburgh	Lock Haven	A	3.0	3.9	4.0	3	6
		B	3.0	3.9	3.0	3	6
		C	3.2	3.0	3.0	3	6
		D	2.8	2.4	2.0	2	5
	Montevideo	A	3.2	2.8	3.0	2	6

Table 6. Factor analysis of five implementation measures.

	ROTATED FACTOR LOADINGS		
Measure	1	2	Communality
Teacher perception	0.052	-0.949	0.902
Sponsor rating	0.864	-0.076	0.752
Consultant checklist	0.670	0.386	0.598
Sponsor checklist	0.888	0.073	0.794
Sponsor overall rating	0.890	-0.184	0.827
Sum of squares	2.779	1.094	3.873

09045

the correlations and the factor analysis: on the subsample we used, the three sponsor measures and the consultant checklist tend to agree on which teachers meet model specifications more than others. The teacher perceptions, however, are tapping something different.

20046

CONCLUSIONS

The analyses in this report generally corroborate the conclusions drawn in the 1970-71 report. In that study we found that models were neither well-specified nor uniformly implemented during the second year of HSPV and that as a result, the interpretation of treatment effects was difficult. We also concluded that because these deviations from the classical experimental ideal seemed to result from the nature of the treatments and the implementation process, they will tend to re-occur in future studies of this kind. If the traditional standards can never be met, they would seem to be inappropriate and an alternative standard is needed for evaluating whether an experimental treatment as implemented is close enough to the ideal that it can be considered to provide a valid test of the treatments. We argued that the key to such a standard is a new definition of full implementation. We recommended, on the one hand, that work should continue to specify all models in as much detail as possible and preferably operational terms. The more completely a treatment is specified, the more likely we are to find that judgments about its implementation can be made on the basis of direct observation and that the criteria for judgments will be made explicit. Both of these factors are important to reducing the subjectivity of the measure and anchoring the response scale. On the other hand, the goal of identical replication should not be pursued but should be shifted to a standard which allows for reasonable variation in the way a model is implemented. At

the same time, some limits must be set on the variation to enable us to distinguish between full implementation, partial implementation and no implementation.

In the present study, with four new implementation measures, we did move a step ahead in specifying the treatments, but as one would expect given the nature of the treatments, not all models were fully specified in operational terms; many checklists are still made up of fairly broad goals and general principles. Consequently, ratings continue to be subjective and therefore are not always easily observable. More important, no alternative standards were set which allowed us to determine whether a class was enough like the model to consider it to be a valid example of the treatment.

As a result, all we can say about implementation again this year is that there is variation among teachers in the extent to which they implement the models to which they were assigned. This pattern was found on all measures, though on some the variation was smaller than on others because of restricted response scales. While we can conclude that all raters distinguish among teachers, and that treatments are therefore not uniform, we cannot make any statements about the size of the variations.

On the positive side, the moderately high agreement among the implementation measures indicates that the measures we have are not without value. Although the scales and model definitions are not refined to the extent that two different raters will assign the same number to a classroom or a teacher's performance,

they do tend to agree on rankings; they know that one teacher is doing a better job of implementing a model than another. If the measures continue to be only ordinal scales, however, then comparisons among classes or sites can safely be made only when they have been rated by the same person. The numbers themselves will continue to have limited meaning. We are more convinced than ever, then, of the importance of establishing standards for assessing implementation within an experimental context. We are not certain, however, that that effort will be successful.

Other conclusions drawn from the work reported here and from our experience with HSPV in general are related to data collection and the development of measures. One major limitation of our analyses has been that they did not include all models. Not only did this have the obvious effect of ruling out conclusions about implementation of those models but it also restricted the comparison of different implementation measures and thus reduced our ability to evaluate the quality of those instruments. Data were missing or excluded for a variety of reasons each of thich suggests a somewhat different remedy.

REC was excluded because it worked with only one site. For the present study this meant we did not have enough data points to evaluate the implementation measures; since there was only one rater, for example, no inter-rater reliability tests could be done. More important, for the experiment in general, a model with a single site means that model effects are always completely

confounded with site effects. The recommendation here is simple (and certainly not new to experimental design): include only treatments with more than one replication in any experiment.

The University of Florida model was excluded from several analyses because the implementation measures focused on different people in evaluating the treatment. The Florida model, unlike the others, is not a classroom program but a home-based program which is implemented by parent educators. Unfortunately some of the implementation measures focused on the parent educators while the others focused on the teachers, making the two sets non-comparable. As with REC, these problems are indicative of a larger issue which runs throughout HPSV: having a treatment of a different type than the other but not taking the differences into account decreases the value of including that treatment. In future experiments a clear position must be taken: either include only treatments which all vary on the same dimension or make plans for adequately assessing a variety of treatments on their own terms. While the second alternative may be the more interesting and productive of the two, it also requires substantially more work than the first.

There were no data on the Enabler model on some implementation measures for two reasons. First, there was confusion throughout HSPV as to whether the Enablers were to be considered as sponsors or OCD consultants. Although they acted as sponsors for this model and, in fact, provided the treatment, they were hired

under the same contractual arrangements as the consultants who monitored implementation. As a result of this confusion, some of the instruments for this model were never administered. Second, and more important, the Enabler model was a new model. The Enablers were not a unified group and never reached the point of drawing up a common model description, thus excluding several other instruments. This experience suggests that a model should not be included in an experiment until it has reached a level of development which allows it to meet certain evaluation requirements such as being able to describe the model program.

Finally, data were missing on some of the implementation measures for the EDC and the University of Kansas models because these sponsors objected to taking the role of evaluator. Both groups feel that their teaching relationship with the site staff is based on a trust which would be violated if they "graded" the staff on their performance. They are not opposed to evaluation per se -- if it is done by someone else, or if it is done within their own format. The Kansas program, for example, has developed its own elaborate system for feedback to the sites and shared their data from that system with us. Since it was not comparable to our other data, however, it was not used in the formal analyses. Regardless of the justifiability of a sponsor not wanting to undermine his or her trustworthiness, it is a problem when participants refuse to cooperate in the evaluation of a project. Perceived conflicts of interest and responsibility to the evaluators should be negotiated and explicitly agreed to in a contract before a study begins.

Such a contractual agreement should also place a responsibility on the evaluators to specify in advance what work they will need from treatment sponsors. Another difficulty in HSPV was that sponsors were deluged with requests for information, for completing forms, and for cooperation in the development of instruments. Not only were these requests numerous, but on occasion the material which resulted was never used. This problem is more general than the two sponsors who did not complete all measures. We suspect that this approach of numerous, repeated, and sometimes ill-conceived requests led all sponsors to be less than conscientious when filling out implementation forms and that probably this compounded the other difficulties with the implementation measures. If roles and obligations could be clarified in advance through an explicit and binding agreement, both sponsors and evaluators should benefit.

Another set of problems in HSPV resulted from the inadequate planning for the design and use of the implementation measures. One such problem was that none of the instruments were field-tested before their use in HSPV . If they had been, a number of problems which affected all the data, such as unreliable response scales, might have been corrected. Since pretesting is expensive and time consuming, especially with treatment-specific measures, it may not be practical to do with all treatments. In HSPV it would have been helpful, however, to request one or two sponsors to be pre-test sponsors and to give them additional

funding to work with the evaluators, and perhaps the consultant-observers, in reliability and validity studies of a variety of instruments.

Another problem in using the implementation instruments, especially the consultant checklists, was that the raters were not systematically trained. Particularly when using an instrument which requires judgments and decisions rather than automatic responses, training is important in providing the raters with common standards of implementation and methods of observation. Its omission before the use of the consultant checklists seriously limited the value of the data collected. For example, we could not determine whether differences among sites in their rated level of implementation reflected real differences or simply differences in the observers.

The key to both these last issues is the recognition that they are important enough to spend time and money on. Unless these steps are scheduled into the data collection process and unless these tasks are made an important and legitimate part of job descriptions, either by having instruments developed prior to an experiment or under a special organizational component, they will continue to fall by the wayside and the data obtained with those instruments will continue to be of limited use.

Perhaps our most general learning in HSPV has been the recognition that some model programs fit into an experimental framework more easily than do others. As we have said, in a classical experiment, it is expected that the treatments to be

tested are specified in detail and that all subjects assigned to a treatment will receive essentially identical replications of it. In HSPV, only the behaviorist models came close to meeting these expectations. The behaviorist models prescribe specific classroom activities of learning more than other models and therefore, classes within those models tend to be more like one another on dimensions related to the models than do classes within other models. Similarly, the implementation of these models is more readily measured because the presence of their components can be more easily observed within a limited time. This is not to say that the behavior models are necessarily superior on theoretical grounds, only that they are more amenable to the experimental situation. The models based on general principles which encourage individual adaptation are more difficult to fit to the experimental mold and their assessment seems to suffer as a result.

Thus, in the study of implementation in the third year of HSPV has revealed that the implementation measures were improved but still not without substantial problems. There are some models which may never fit into a quanitfied, easily observable measure. In those models, and in others like them, there may always be an element of subjectivity in assessing the extent to which a program is implemented. The question of implementation, however,

10651

is a crucial one: determining the presence of a treatment is a necessary prerequisite for meaningfully assessing program effectiveness. We continue to believe, then, in the importance of establishing standards for evaluating the adequacy of implementation for experimental purposes. We also recommend that time and money be set aside for the development of implementation measures in any social experiment or program evaluation. Finally, we recommend that a formal agreement be drawn up between evaluators and people responsible for treatment implementation which outlines their responsibilities to each other.

30055

Appendix A

ACCOUNTING FOR SITE TO SITE VARIATION IN LEVELS OF IMPLEMENTATION

In the 1970-71 implementation report, we presented a regression equation in which three site characteristics accounted for over 97% of the variation in levels of implementation as measured by sponsor ratings of teachers on a 0-9 scale. The site characteristics were taken from a form on which sponsors were asked to rate sites on the frequency or extent of occurrence of seventeen variables. The three which entered the equation were intra-staff friction, adequacy of physical plant indoors, and rapport between administrators and staff. Table A-1 shows this equation. We said at that time that at least two interpretations of the findings were possible. On the one hand, there may be a real relationship between the pleasantness of a site and implementation: pleasantness may be an indication that a site is operating smoothly and as such be a prerequisite for implementation. On the other hand, the finding may only indicate a failure to distinguish between good implementation and a generally nice place.

With an equation which accounts for such a large proportion of the variance, replication is important. Exact replication -- using identical variables with new data -- was impossible because the form from which the sponsors' judgments about site characteristics were taken was not administered during the third year of the

Table A-1

Regression Analysis with 1970-71 Data

Variable Name	b	SE_b	b*	T-Test	df	Significance
Intra-staff friction	-0.5276	0.027	-0.658	-19.80	23	$< .001$
Rapport between administration and staff	0.4232	0.031	0.466	13.50	23	$< .001$
Adequacy of physical plant indoors	0.2025	0.025	0.265	8.05	23	$< .001$
Regression Constant	4.573					

R^2 = 0.977
R = 0.988
SD_{res} = 0.161

F = 326.96 with 3 and 23 degrees of freedom ($p < .001$)

Note: This is a stepwise regression: variables were allowed to enter in order of importance

b: regression coefficient
b^*: standardized regression coefficient
R^2: multiple correlation coefficient squared

study. This decision was made before the second year data were analyzed when we did not anticipate the importance of the information. The equation was approximated by creating variables three items from the evaluation reports submitted by the Head Start directors:

60. Have any of the following groups at your centers had disagreements during the year? (Check any which have occurred)
___ teachers and aides
___ professionals and paraprofessionals
___ teaching and supervisory staff
___ ethnic groups
___ administrative staff and parents
___ intra-staff

1=check
0=blank

62. Has any of the staff conflict been a serious obstacle to a good program?

3=very serious
2=moderately serious
1=trivial

65. Mark the following list of equipment and materials to show which you have or need:

Outdoor playing facilities like slides and jungle gyms
Cots or beds
Toy or household equipment, cooking utensils, dramatic play, clothing
Nature or science materials
Building materials
Printing devices
Art materials
Musical instruments
Audio-visual equipment
Special mechanical or electronic teaching devices
Reading material related to curriculum
Books about children with background similar to pupils

1=have
2=need
3=don't have or need
4=have some, need more or 1,2 in same space

A variable labelled "groups having conflict" and with values between 1 and 6 was created from item 60 by simply summing the number of items checked such that a higher score indicates that more groups have had conflict. Item 62 was taken as it is to create a variable labelled "conflict is obstacle"; again, a higher score indicates more conflict. A third variable, "adequacy of materials and equipment," was formed from item 65 by eliminating the 3 and 4 responses because they are qualitatively different and averaging the 1 and 2's; the resulting variable scores range between 1 and 2 with a higher score indicating higher need for additional materials.

Table A-2 presents the analysis which uses these variables to account for the variance in levels of implementation, again as measured by sponsors rating teacher performance on a ten-point scale. As can be seen, the equation does not even yield a significant R^2 (amount of variance explained). Admittedly, the variables are extremely rough and the roughness may in part explain the failure to account for a large portion of the variance. But if the equation from the second year analysis represents a real relationship between site pleasantness and implementation, that relationship should show up to some extent even with a different person doing the rating and with somewhat different variables. This finding of almost no relationship suggests that earlier findings may have been a function of the sponsors' equation of implementation with site pleasantness.

39959

Table A-2

Regression Analysis With 1971-72 Data

Variable Name	b	SE_b	b*	T-Test	df	Significance
Adequacy of materials and equipment	1.8734	1.353	0.337	1.38	21	.181
Conflict is obstacle	-0.4859	0.399	-0.294	-1.22	21	.238
Groups having conflict	-0.1265	0.631	-0.042	-0.20	21	over .500
Regression Constant	4.538					

Correlation Square = 0.103
Correlation = 0.321
Deviation of Residuals = 1.062

F = 0.80 with 3 and 21 degrees of freedom (p over .500)

Note: This is a stepwise regression: variables were allowed to enter in order of importance

b: regression coefficient
b^*: standardized regression coefficient
R^2: multiple correlation squared.

Appendix B

1971-72 IMPLEMENTATION MEASURES

This appendix includes:

1. Sample consultant checklists: the checklists from the Bank Street and the Oregon models are included here as examples of the form of the checklists and the types of components which make up two very different programs. Copies of all checklists can be found in Monaghan's report.

2. Sponsor checklists: these are shortened versions of the consultant checklists.

3. Sponsor overall ratings: sent as cover form with the sponsor checklists; the same form was used for all models.

4. Sponsor ratings: sent out by the Stanford Research Institute; the same form was used for all models.

The teacher perception item is presented in the text of the report on page 25.

KEY
1=Specification not at all implemented
2=Specification implemented to some extent
3=Specification implemented to a great extent
4=Fully implemented
5=Beyond technical implementation
X=No opportunity to observe

MODEL IMPLEMENTATION CHECKLIST: 1971-72
BANK STREET (FEBRUARY REVISION)

I. LEARNING-TEACHING BEHAVIOR

A. Children's Behavior

1.	Children move freely but purposefully among many interest areas.	1	2	3	4	X
2.	Children demonstrate active participation in their own learning through self-initiated expression and through seeking more understanding of facts, ideas and processes.	1	2	3	4	X
3.	Children make choices among activities and are able to act upon them.	1	2	3	4	X
4.	Children select materials that are appropriate to the activity.	1	2	3	4	X
5.	Children appear to find their activities satisfying.	1	2	3	4	X
6.	Children show persistence in pursuing a given task.	1	2	3	4	X
7.	Children generally observe agreed upon limits to their behavior.	1	2	3	4	X
8.	Children carry through steps in working with materials, i.e.: taking out, using and putting them away in designated locations.	1	2	3	4	X
9.	Children show some degree of understanding of the identity, characteristics, function, category and relationships of objects and persons.	1	2	3	4	X
10.	Children show interest in probing, discovering, experimenting and analyzing the outcome of their experimentation.	1	2	3	4	X
11.	Children organize their ideas, reason, plan and solve problems.	1	2	3	4	X
12.	Children exercise the freedom to express feelings, spontaneously.	1	2	3	4	X

B. Adult Behavior

*1.	The adult is supportive of exploration and questioning,	a 1 2 3 4 5 X b
2.	The adult provides opportunities for skill development.	a 1 2 3 4 5 X b
3.	The adult stimulates and encourages creativity.	a 1 2 3 4 5 X b
4.	The adult challenges and supports problem solving and coping behavior.	a 1 2 3 4 5 X b
5.	The adult shows respect for children's ideas and helps them clarify and extend their thinking.	a 1 2 3 4 5 X b
6.	Adults show concern for rights and feelings of children.	a 1 2 3 4 5 X b
7.	The adults help children to evaluate their own behavior and its consequences for themselves and others.	a 1 2 3 4 5 X b
8.	The adults plan limits with the children -- limits which are rational, understood, and acceptable.	a 1 2 3 4 5 X b
9.	Inappropriate behavior is not dealt with by censure but is redirected.	a 1 2 3 4 5 X b
*10.	The adults take into account each child's interests, strengths, weaknesses, and learning styles in developing individualized curriculum.	a 1 2 3 4 5 X b
11.	The adults encourage children to work cooperatively and to interact in many ways with one another.	a 1 2 3 4 5 X b
12.	Adults encourage children to describe out-of-school experiences, and show interest in the child's whole life.	a 1 2 3 4 5 X b
13.	Adults expect and acknowledge task completion when the task is appropriate to the child's age and capacity.	a 1 2 3 4 5 X b

*This item may require interview rather than observation alone.

14.	Adults record and place displays of children's language and art work prominently and with respect.	-	1	2	3	4	5	X
		a						
		b						
15.	The adult places displays at child's eye level.	-	1	2	3	4	5	X
		a						
		b						
16.	Adults change displays frequently to reflect children's changing interests.	-	1	2	3	4	5	X
		a						
		b						
17.	Each adult provides a role model with which the children may make positive identification.	-	1	2	3	4	5	X
		a						
		b						
18.	The staff helps to create an open, relaxed, active atmosphere in class.	-	1	2	3	4	5	X
		a						
		b						
19.	Adults encourage children to plan their own work without adult interference but become involved in their activities as needed.	-	1	2	3	4	5	X
		a						
		b						
20.	Adults provide opportunities for child-to-child interaction.	-	1	2	3	4	5	X
		a						
		b						

C. Interaction in Classroom - (Use Key on first page)
(Child-adult, child-child, adult-adult)

1.	Children appear to trust adults.	1	2	3	4	X	
2.	Children interact, cooperate and share ideas and material with other children.	1	2	3	4	X	
3.	A teacher and two assistants share the operational responsibilities of the classroom equally.	1	2	3	4	5	X
3a.	The diagnosis and prescriptive responsibilities are primarily responsibilities of the teacher.	1	2	3	4	5	X
4.	Suggestions from the assistants are sought by the teacher and there is joint planning of each day's activities.	1	2	3	4	5	X
*5.	There is continuing interaction, sharing of information and insights, and mutuality of goals between teaching staff and ancillary staff (such as nurses, family workers and guidance personnel) with the latter observing and sometimes participating.	1	2	3	4	5	X

*This item may require interview rather than observation alone.

*6.	Parents interact with children and adults in terms of children's learning, as the parents participate in the classroom and on trips.	1	2	3	4	X	
*7.	There is one-to-one conferencing between parents and teachers concerning the learning needs of individual children. (Interaction with other local staff and the community is described under the section on sponsor's delivery system.	1	2	3	4	X	
II.	LEARNING ENVIRONMENT						
A.	Organizational						
1.	The schedule includes work and play periods, both indoors and outdoors, naps, snacks, and luncheons, which are all treated as active learning periods.	1	2	3	4	X	
2.	Field trips occur once or twice a week, usually within the school or the neighborhood.	1	2	3	4	X	
3.	Children's groupings are flexible so as to provide maximum opportunity for choice and mobility, with the exception of a few formal, structured groups, such as meeting time (or circle time) and transition time. Meeting time may be for the whole group or a small group.	1	2	3	4	X	
4.	Most activities, whether formal or informal, are carried out in small groups or individually.	1	2	3	4	X	
5.	Teachers and aides move from group to group during the day as needed.	1	2	3	4	5	X
6.	Lunch is served family style, i.e., in small groups, with the children serving themselves.	1	2	3	4	X	
6a.	During lunch, adults stress child-child and child-adult conversation.	1	2	3	4	5	X
B.	Physical	a					
		b					
1.	The classroom is structured and ordered to provide defined areas within which children work and play with materials appropriate for each interest center.	1	2	3	4	X	

*This item may require interview rather than observation alone.

2.	Materials are located so that children know where they are and can reach them readily.	1	2	3	4	X
3.	Materials and equipment are planned, rotated, replenished and changed as the need arises.	1	2	3	4	X
4.	There is emphasis upon use of natural materials within the child's own environment, and child-made, teacher-made and parent-made materials and equipment, as well as commercial items.	1	2	3	4	X
5.	Materials are appropriate to the age level. such as including large blocks for motor activity and water, sand and clay for sensory activities.	1	2	3	4	X
6.	The interest areas are separated from the main activity of traffic.	1	2	3	4	X
7.	There is a protected area with a minimum of interruption for a child or group desiring to engage in a quiet activity.	1	2	3	4	X
8.	There is private space for each child's own things.	1	2	3	4	X
9.	Adequate storage space is available and storage equipment is often used as room dividers.	1	2	3	4	X

C. Social-Emotional

1.	Children convey a sense of joy as they work in the room.	1	2	3	4	X	
2.	There is awareness of children's feelings, in a way of the primary causes for doing or not doing things with one another.	1	2	3	4	X	
3.	Adults express their own feelings, in a way that is natural but also geared to children's understanding.	- 1	2	3	4	5	X
		a _____					
		b _____					
4.	Most children show sympathy for other children when they are in trouble, rather than ridicule or teasing.	1	2	3	4	X	
5.	Most children show friendliness and warmth toward other children.	1	2	3	4	X	
6.	Most children seem relaxed and comfortable instead of tense, suspicious or fearful.	1	2	3	4	X	

7. Children express feelings of approval or disapproval openly, naturally in a milieu of acceptance.	1	2	3	4	X	
8. Children's efforts at inner control and mastery of their own world are recognized and encouraged.	1	2	3	4	X	

III. CURRICULUM

A. Basic Principles Undergirding Specific Activities*

1. Curriculum is differentiated and particularized, i.e. it reflects the individual goals, interests and needs of children.	1	2	3	4	5	X
2. Curriculum is structured according to basic educational principles but is completely flexible in response to the developmental stages of the children, their evolving competencies, and opportunities for learning as they arise in each situation.	1	2	3	4	5	X
3. Curriculum moves from the concrete to the conceptual, through first-hand experiences the meaning of which is clarified for and with the children.	1	2	3	4	5	X
4. The curriculum is based upon the adult's study of how each child organizes and reinterprets his experience through "play" and his own choice of activities.	1	2	3	4	5	X
5. Curriculum is relevant to life styles of families and community.	1	2	3	4	5	X
6. The curriculum reinforces and extends the adult's supportive interaction with children aimed to develop positive feelings about self.	1	2	3	4	5	X
7. Currriculum planning is a team operation with the teacher primarily responsible and accountable for its effectiveness but drawing upon the knowledge and insights of both paraprofessionals and ancillary staff.	1	2	3	4	5	X
8. Curriculum is directed toward cognitive and affective growth in constant interaction.	1	2	3	4	5	X
9. Curriculum provides for motor and sensory experiences for independent investigation for problem-solving and for the development of concepts such as sequential ordering, symbolic	1	2	3	4	5	X

*This whole category may require interview rather than observation alone.

9. (Cont)representation, categorizing, spatial and time relationships, identifying characteristics, functions,and a wide variety of relationships among objects and persons. (The results of these curriculum emphases are delineated under the section on Child Behavior I, 17.

10. The activities and materials are centered around various themes such as: home (cooking, household chores); environmental studies (rocks, soil, plants, etc.); creative expression (blocks, dramatic play, music, art, etc.). 1 2 3 4 5 X

B. Language

1. Language, both written and spoken, surrounds the child throughout work and play periods. 1 2 3 4 5 X

2. Language activities include stories read to and by children, experiences dictated and illustrated by children, experience charts, name cards, labelling and frequent free discussion periods. 1 2 3 4 5 X

3. Resources for such activities are drawn from books, pictures, films, and various other media as well as the children's own experiences at home, in school and in organized field trips. 1 2 3 4 5 X

4. Language is learned as a useful pleasurable tool. 1 2 3 4 5 X

5. The curriculum develops not only basic skills but also competencies at a higher level including the creative use of language in self-expression. 1 2 3 4 5 X

6. A wide variety of commercial materials are creatively adapted to the Bank Street Approach. 1 2 3 4 5 X

C. Math

1. Math also pervades the school day and numerical concepts are related functionally to everything that happens in the classroom. 1 2 3 4 5 X

2. Math is learned through planned activities which are clearly applicable to the "here and now" of the child's world. 1 2 3 4 5 X

3. Math activities include use of manipulative materials such as cuisenaire rods, differentiated by age level and individual competencies. 1 2 3 4 5 X

4.	There is continuing practical application in such areas as cooking, block building, field trips, taking attendance, measuring, weighing, care of animals, and in relation to time and space.	1	2	3	4	5	X
5.	Specific skills are learned sequentially while gradually exposing the child to more complex concepts and mathematical experiences.	1	2	3	4	5	X

D. Environmental Studies

1.	The curriculum moves from themes of home and school into exploration of the neighborhood and finally the broader environment, in terms of both time and space.	1	2	3	4	5	X
2.	As in other aspects of the curriculum there is a constant tuning in to the child's interest and building with him environmental studies that will seem both important and pleasurable to him.	1	2	3	4	5	X
3.	Trips to the immediate community include exploration of food, housing, transportation, public services, and particularly the people who provide such services.	1	2	3	4	5	X
4.	Environmental studies include both the physical and social sciences as they relate to one another and also in their discrete aspects.	1	2	3	4	5	X
5.	Studies of the physical environment are rich in raw materials, such as water, wind, ice, plants, animals and in experimentation with scientific procedure such as growing plants, feeding animals, and the transformation of ingredients during the cooking process.	1	2	3	4	5	X

E. Art, Music, Dramatic Play

1.	Creative experiences are both planned and spontaneous and like other aspects of the curriculum are integrated into the totality of the child's learning.	1	2	3	4	5	X
2.	Musical activities include rhythm, use of musical instruments, singing (structured or spontaneous) and listening to music in various forms as played by staff, parents and invited guests as well as through records and tapes.	1	2	3	4	5	X

3. Art is also both spontaneously expressive through finger painting, murals, easel painting, and interpretive of art created by others as seen in books, pictures, museums, displays.	1	2	3	4	5	X
4. Dramatic play is viewed not only as a valued outlet for the child himself but also as one of the most important diagnostic tools for the adults' insights into children's genuine feelings, anxieties, strengths, and potential for development.	1	2	3	4	5	X

IV. SPONSOR'S DELIVERY SYSTEM *

A. Commitments of Sponsor and Community

1. The Sponsor has developed a theory on Early Childhood Education. The Sponsor has and is developing a practical program for individualization which involves interpretation, stimulation, and staff development to enable the community to put the approach into practice.	1	2	3	4	X
2. The contractual relationship assumes that the community has a genuine interest in the primary schools serving a population of poor children and that it will support the efforts of the sponsor adminstratively.	1	2	3	4	X

B. Interpretation and Assistance in Implementation

1. The Director of the Bank Street program maintains a continuing relationship as both consultant and catalyst in all areas of the program: administrative, educational, supportive services, parent involvement, and community relations.	1	2	3	4	X
2. An Associate Director coordinates Program Analysis and another Associate Director gives priority to coordinating the Head Start Program.	1	2	3	4	X
3. The Sponsor assigns a continuing field representative to each project who not only interprets the sponsor's approach but also works with local staff and parents in its implementation.	1	2	3	4	X
4. The Sponsor provides specialists in specific areas as resource persons on-site and at the College.	1	2	3	4	X

*All categories under Section IV may require interview rather than observation alone.

5. The Sponsor provides interpretive materials such as working papers, curriculum guides, self-study forms, questionnaires, reports, staff evaluational forms, and various media such as films, carousels, video-tapes, as well as Bank Street published materials including basal readers, discovery materials and language stimulation materials. 1 2 3 4 X

C. Staff Development

1. The Sponsor conducts institutes and workshops at Bank Street for Directors, principals and other administrative personnel in the system, staff developers, teaching teams, parents, ancillary staff. 1 2 3 4 X

These institutes include observation of the Demonstration School for Children, visits to N.Y.C. Follow Through program in P.S. 243, and other educational experiments in the New York area as well as seminars, individual consultations and team planning.

2. The Sponsor assists local staff in developing on-site institutes and workshops covering the same type of personnel but including all local individuals in the respective categories. These institutes are held before the school year begins and throughout the year as needed. They cover educational support for each component, assistance in diagnostic teaching with emphasis on psychological, social and parent liaison services as supportive of teaching, and the development of new ways of assessing individual child growth along many dimensions. 1 2 3 4 X

3. There is continuing emphasis upon supporting and extending team work throughout both institutes and consulting services including team work: 1 2 3 4 X
   a) between administrative and program personnel.
   b) between field representatives and staff developers.
   c) between staff developers and teaching teams.
   d) within the teaching teams
   e) between all of the above and ancillary staff.
   f) with and among parents

4. The consultant services of field representative, resource persons and central Bank Street staff are viewed as an essential component of the circular process of staff development. The Sponsor provides input to the community, considers feedback from the community and eventually plans jointly with the community to meet differentiated needs with continuing support and guidance. 1 2 3 4 X

5.	The Sponsor organized courses for the teaching personnel and parents with academic credit to further Career Development.	1	2	3	4	X

D. Program Analysis

1.	Teachers are using instruments for systematic classroom observation which are designed not only to assess progress toward implementation of the Bank Street Approach but also to strengthen self analysis and staff development.	1	2	3	4	X

V. PARENT INVOLVEMENT*

A. In Children's Learning

1.	Parents are encouraged to participate in the school's learning activities, such as helping with story time, field trips, cooking, and making materials.	1	2	3	4	X
2.	Parents who participate are invited to conference with the teaching team so as to share thoughts, feelings and ideas about program and children.	1	2	3	4	X
3.	Special skills and talents of parents are searched out and utilized in the classroom, to enrich the curriculum and to enhance positive feelings about self for both parents and children.	1	2	3	4	X
4.	Parent and teacher enter into one-to-one conferencing around the learning needs of each parent's child or children, in which the teacher learns from and with the parent about a specific child.	1	2	3	4	X
5.	Parents and teachers cooperate in planning out-of-school reinforcement of what children are doing and learning in school, which is facilitated by home visits by both the teaching teams and ancillary staff.	1	2	3	4	X
6.	Parents develop and/or use checklist for classroom observation, i.e. what to look for in a Bank Street-sponsored classroom.	1	2	3	4	XX

*All categories under Section V may require interview rather than observation alone.

B. In Decision-making through PC (Policy Committee)

1. Parents are involved in decision-making through the appropriate representation and responsibility in policy-making boards.	1	2	3	4	X
2. The nature and scope of such decisions are determined by the extent to which parents understand and participate in the school's program and also by the stage of organization and functioning which the policy board has reached.	1	2	3	4	X
3. Parents are encouraged and assisted when such assistance is desired, in organizing their policy board, in such matters as by-laws, contracts, committee structure and program planning.	1	2	3	4	X

C. In Self-Development

1. Parents participate in study groups often initiated by parents themselves.	1	2	3	4	X
2. Parents' workshops are designed to further interpretation of the Sponsor's Approach to Early Childhood Education and the school's implementation of the model.	1	2	3	4	X
3. Parents are being trained to interview each other, using the new Questionnaire for Parents.	1	2	3	4	X
4. Parents who have actively involved the above mentioned activities interpret the model and the school's educational goals to other parents.	1	2	3	4	X
5. Parents may enroll in courses for high school equivalency or for undergraduate credit, particularly those who are employed as para-professionals.	1	2	3	4	X

D. In School-Related Social Activities

1. Parents plan and conduct social activities for parents themselves.	1	2	3	4	X
2. Parents assist in planning and conducting children's social activities such as end-of-year functions, birthday parties, holiday festivals, which bring the ethnic culture of the child into the school.	1	2	3	4	X
3. In some communities a special room or house may be set aside for the use of parents, which is often used for educational and community activities as well as for the primary social function.	1	2	3	4	X

E. In Community Action

1. Parents participate in existing Community Action programs which affect directly or indirectly their children's learnings.	1	2	3	4	X
2. In some communities the Head Start parents spearhead new programs and create new organizations.	1	2	3	4	X
3. Head Start parents cooperate with Follow Through parents in the above mentioned activities.	1	2	3	4	X
4. Parents interpret the program not only to other parents but also to the broader community.	1	2	3	4	X

MODEL IMPLEMENTATION CHECKLIST: 1971-72
OREGON

KEY 1=Specification not at all implemented
2=Specification implemented to some extent
3=Specification implemented to a great extent
4=Fully implemented
5=Beyond technical implementat implementation

I. MATERIALS AND ORGANIZATION

A. Teachers and aides are following a schedule including 1 1/2 hours devoted to basic skill using Distar materials, time for snack, lunch, hygiene, recess, and art activities.	1	2	3	4	5
B. Each teacher or aide teaches each child one lesson each day in the reading, arithmetic and language parts of the Distar materials.	1	2	3	4	5
C. Teachers and aides, are beginning and ending Distar lessons on time.	1	2	3	4	5
D. Teachers and aides are assigned permanently to specific instructional areas.	1	2	3	4	5
E. One fourth of the day is alloted to children's self-selection of activities.	1	2	3	4	5
F. Three or more instructional groups of 4-9 children have been formed, on the basis of pre-testing by teacher and aides.	1	2	3	4	5
G. Records are kept of the continuous progress of each group in each subject.	1	2	3	4	5
H. Provision is made for tutoring any child requiring additional instruction on any subject by the teacher or aide who teaches that subject.	1	2	3	4	5
I. Teachers and aides effectively utilize "free time" when children enter in the morning to support language learning.	1	2	3	4	5
J. Teachers and aides participate in weekly in-service meetings.	1	2	3	4	5

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Norma Bernstein-Tarrow The Huron Inst.
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

© Copyright, 1971: Anne-Coolidge-Monaghan/ The Huron Institute

Number of school days since last consultant site visit ___

	ARITHMETIC			READING			LANGUAGE		
	Top Group	Middle Group	Low Group	Top Group	Middle Group	Low Group	Top Group	Middle Group	Low Group
Day in program (page number) of group on last visit									
Day in Program (page number of group on present visit									
Difference in page numbers between last and present visits									

II. FORMAL INSTRUCTIONAL PERIODS

A. Teacher[1] Preparation and Presentation

1. Teachers know format of lesson and look down at book only for examples.	___		1	2	3	4	5
	___	a	1	2	3	4	5
	___	b	1	2	3	4	5
2. Teachers are properly using signals to get attention and insure responding by all children at once.	___		1	2	3	4	5
	___	a	1	2	3	4	5
	___	b	1	2	3	4	5
3. Teachers use correct terminology.	___		1	2	3	4	5
	___	a	1	2	3	4	5
	___	b	1	2	3	4	5
4. Teachers do not vary from format unless otherwise specified by the consultant or curriculum supervisor.	___		1	2	3	4	5
	___	a	1	2	3	4	5
	___	b	1	2	3	4	5
*5. Teachers demonstrate ability to extend concepts through the use of their exemplars and concrete materials.	___		1	2	3	4	5
	___	a	1	2	3	4	5
	___	b	1	2	3	4	5
6. Teachers vary pacing to suit the needs of the children or the situation.	___		1	2	3	4	5
	___	a	1	2	3	4	5
	___	b	1	2	3	4	5
*7. Teachers are properly diagnosing causes of error and employing appropriate correction paradigms.	___		1	2	3	4	5
	___	a	1	2	3	4	5
	___	b	1	2	3	4	5
8. Teachers have all materials on hand and clearly marked to assure continuous flow of lesson.	___		1	2	3	4	5
	___	a	1	2	3	4	5
	___	b	1	2	3	4	5

1. Teacher means any permanent, paid adult staff member in a class.

* Please briefly note anecdotal evidence with which rating is made underneath starred item.

*9. Teachers are adequately employing individual tasks.			1	2	3	4	5
		a	1	2	3	4	5
		b	1	2	3	4	5
10. Provision is made for the completion of Take-Homes in each subject area during small-group instruction and distribution of Take-Homes to the children before they are dismissed.			1	2	3	4	5
		a	1	2	3	4	5
		b	1	2	3	4	5
11. Teachers are accomplishing criterion learning within 3-6 days on all new motor skills and within 2 days on all other new skills.			1	2	3	4	5
		a	1	2	3	4	5
		b	1	2	3	4	5
B. Control and Reinforcement							
1. Each teacher uses systematic, positive reinforcement whenever possible in working with children.			1	2	3	4	5
		a	1	2	3	4	5
		b	1	2	3	4	5
2. Teacher is consistent in ignoring off-task behavior as much as possible.			1	2	3	4	5
		a	1	2	3	4	5
		b	1	2	3	4	5
3. Whenever necessary, teachers are able to apply systematic behavior modification punishment techniques.			1	2	3	4	5
		a	1	2	3	4	5
		b	1	2	3	4	5

III. REMAINDER OF PROGRAM (Unspecified Time)

Please note by underlining which of the following activities occur and comment on how they are carried out below.

A. Creative activities

B. Field trips

C. Provision for physical-motor skills development

D. Children's taking responsibility, planning or evaluation

E. Children self-selection of activities

F. Provision for stories, dramatizations, role playing, etc.

G. Provision for class discussions, sharing, etc.

Equipment present:

Language-Library _____
Creative _____
Science _____
Number _____
Building _____
Housekeeping _____
Other _____

SPONSOR IMPLEMENTATION MEASURE

MODEL: FAR WEST

SITE:

KEY: 1 = Specification not at all implemented.
2 = Specification implemented to some extent.
3 = Specification implemented to a great extent.
4 = Technically implemented.

You are rating the class with the following personnel, please circle, beneath each teacher's name, the year in which he or she is now teaching in the model.

Teacher 1: 1st 2nd 3rd	Teacher 2: 1st 2nd 3rd	Teacher 3: 1st 2nd 3rd
1. The room includes both suggested arrangement of space and model equipment.		Classroom: 1 2 3 4
2. Teachers and aides work together with no differential responsibilities during class time, especially in regard to teaching-learning duties, spending over half their time with individual children or small groups.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
3. Early in the program, the teachers direct work toward basic concepts of color, position and relation. The remainder of the program stresses self-rewarding experiences and problem solving activities.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
4. The teacher observes children, letting them take the initiative, supporting them in their work, responding verbally to children and asking questions that require more than a yes or no answer.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
5. Teachers emphasize positive redirection as major technique of handling inappropriate behavior.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
6. Children move at their own pace, exploring materials chosen by the children 75% of the time and which they do for the internal satisfaction of doing them.		Children: 1 2 3 4

FAR WEST PAGE TWO

SITE:

7. In the specific area of problem solving, children are solving problems (physical and personal), are taking reasonable risks of failure, and maximizing their resources. Children: 1 2 3 4

8. Children cooperate and interact with both peers and adults, requiring little external control and cooperating well with their own emotions. Children: 1 2 3 4

I am very familiar _____ with this class.

I am somewhat familiar _____ with this class.

I am not too familiar _____ with this class.

Signed _____

Date _____

SPONSOR IMPLEMENTATION MEASURE

MODEL: ARIZONA

SITE:

KEY: 1 = Specification not at all implemented.
2 = Specification implemented to some extent.
3 = Specification implemented to a great extent.
4 = Technically implemented.

You are rating the class with the following personnel, please circle, beneath each teacher's name, the year in which he or she is now teaching in the model.

	Teacher 1: 1st 2nd 3rd	Teacher 2: 1st 2nd 3rd	Teacher 3: 1st 2nd 3rd
1.	The room is divided into interest centers; there is an expectation for each child to be responsible for his involvement in a center.		Classroom: 1 2 3 4
2.	There are many materials that invite exploration and manipulation as well as skill development, such as intellectual kits; there is a rich reading environment; the materials are attractive and easily accessible, and there is a place for children to store their own personal materials.		Classroom: 1 2 3 4
3.	Teaching is orchestrated; the teacher uses a variety of methods to foster the development of intellectual skills, and the teacher and the aide mediate the environment as children interact with the environmental activities by explaining reasons for actions and through frequent use of open-ended questions.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
4.	The teacher and aide continually try to develop children's language by accepting the child's language, and by modeling through "corrective feedback" wherever applicable and expanding the child's language whenever possible; the teacher uses the child's language in teaching learning skills.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
5.	The teacher provides, when possible, options and encourages pupils to make choices, by having a wide range of materials available, by verbally posing selections for the children, and by displaying an attitude that the choices are indeed acceptable.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4

ARIZONA

SITE:

6. The teacher uses reinforcement to bring about behavior change: she reinforces verbally and by means of arranging the environment; she provides relevant success oriented activities; she helps children accept necessary social restraints; she praises and identifies the behavior that is being singled out for praise; she considers herself a model and realizes that much learning will be gained by imitation.

Teacher 1: 1 2 3 4
Teacher 2: 1 2 3 4
Teacher 3: 1 2 3 4

7. The teacher individualizes her teaching on the basis of her records and her general assessment of the needs of children; she understands pupil abilities and different levels of performance; she accommodates individual needs by providing an abundance of materials that develop the same skill.

Teacher 1: 1 2 3 4
Teacher 2: 1 2 3 4
Teacher 3: 1 2 3 4

I am very familiar _____ with this class.

I am somewhat familiar _____ with this class.

I am not too familiar _____ with this class.

Signed _____

Date _____

SPONSOR IMPLEMENTATION MEASURE

MODEL: Bank Street

SITE:

KEY: 1 = Specification not at all implemented.
2 = Specification implemented to some extent.
3 = Specification implemented to a great extent.
4 = Technically implemented.

You are rating the class with the following personnel, please circle, beneath each teacher's name, the year in which he or she is now teaching in the model.

Teacher 1: 1st 2nd 3rd	Teacher 2: 1st 2nd 3rd	Teacher 3: 1st 2nd 3rd
1. The room is composed of interest areas with some quiet space for children, a private place for each child's own things and adequate available general storage.		Classroom: 1 2 3 4
2. Materials both "home-made" and commercial, are appropriate to age level of children, are located properly and reflect changing needs of the class.		Classroom: 1 2 3 4
3. The schedule, routine and groupings are those which the model specify.		Classroom: 1 2 3 4
4. The general principles of curriculum are to direct specific activities toward both cognitive and affective growth, covering concepts (e.g., representation, categorizing) as well as language, math, environmental studies, art, music, and dramatic plays. Curriculum is relevant to each child as well as his or her larger community.		Classroom: 1 2 3 4
5. The adults both support and provide for the intellectual skills which the model specifies for children.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
6. The adults show concern for the rights and feelings of children, helping them evaluate their behavior and planning limits with them.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
7. The adults provide opportunities for good relationships between children and adults as well as between children and children.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4

Bank Street PAGE TWO

SITE:

8. Adults individualize curriculum, expecting and acknowledging task completion and, in general, valuing the work of children.	Teacher 1:	1 2 3 4
	Teacher 2:	1 2 3 4
	Teacher 3:	1 2 3 4
9. The staff interacts and works together according to specifications of the model.	Teacher 1:	1 2 3 4
	Teacher 2:	1 2 3 4
	Teacher 3:	1 2 3 4
10. Parents are involved either in that part of the program which revolves around children, participate in PAC according to the model, or participate in activities other than those directly related to the classroom.	Parents:	1 2 3 4
11. Children carry through all steps of work in relation to materials, choose among activities, moving among them freely and purposefully and enjoying their work.	Children:	1 2 3 4
12. Children display desirable thinking processes of the sort which the model is attempting to promote.	Children:	1 2 3 4

I am very familiar _____ with this class.

I am somewhat familiar _____ with this class.

I am not too familiar _____ with this class.

Signed _____

Date _____

SPONSOR IMPLEMENTATION MEASURE

MODEL: OREGON

SITE:

KEY: 1 = Specification not at all implemented.
2 = Specification implemented to some extent.
3 = Specification implemented to a great extent.
4 = Technically implemented.

You are rating the class with the following personnel, please circle, beneath each teacher's name, the year in which he or she is now teaching in the model.

Teacher 1: 1st 2nd 3rd	Teacher 2: 1st 2nd 3rd	Teacher 3: 1st 2nd 3rd
1. The part of the schedule devoted to the model is 1-1/2 hours, with each teacher assigned permanently to an instructional area, teaching small groups of 4-9 children or less in Distar materials: reading, arithmetic or language, beginning and ending these lessons on time.		Classroom: 1 2 3 4
2. 1/4 of the day is allotted to children's self-selection of activities.		Classroom: 1 2 3 4
3. Teacher presentation of lessons follows model specifications by use of correct terminology, proper use of signals, and use of appropriate correction paradigms.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
4. Teachers appropriately respond to individual variations of students by changing pace of lessons, use of their examples and concrete material, and employment of individual tasks.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
5. Teachers use of control and reinforcement follows model specifications.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4

I am very familiar _____ with this class.
I am somewhat familiar _____ with this class.
I am not too familiar _____ with this class.

Signed _____

Date _____

SPONSOR IMPLEMENTATION MEASURE

MODEL: KANSAS

SITE:

KEY: 1 = Specification not at all implemented.
2 = Specification implemented to some extent.
3 = Specification implemented to a great extent.
4 = Technically implemented.

You are rating the class with the following personnel, please circle, beneath each teacher's name, the year in which he or she is now teaching in the model.

Teacher 1: 1st 2nd 3rd	Teacher 2: 1st 2nd 3rd	Teacher 3: 1st 2nd 3rd
1. The curricular areas of math, reading and handwriting are taught with prescribed materials (though many acceptable materials fall under the "readiness" category).		Classroom: 1 2 3 4
2. The daily schedule can alternate between 2-4 earn periods and 2-4 exchange periods; around 15% of the day is devoted to academic areas (though the range of acceptance is from 15-30%); work periods increase in length over the year from approximately 10 minutes at the beginning to a range of 15-40 minutes; exchange periods vary from 10 to 45 minutes; there is outdoor play at least once a day and music/dramatic play at least once a week.		Classroom: 1 2 3 4
3. Teachers use and understand positive reinforcement and the token economy as prescribed by the model.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
4. Teachers interact with the children as much as possible during the spend time by playing or working with the children.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
5. All children can have experience in each of the three basic curriculum areas once a day during "earn" periods in groups of 4-6 or less; tutoring is almost always 1 to 1.		Children: 1 2 3 4

KANSAS PAGE TWO

SITE:

6. Parents receive an orientation program at the beginning of the year and before each rotation of the parents; parents are rotated at 6 week intervals in each class; parents are trained by a parent trainer.

Parents: 1 2 3 4

I am very familiar _____ with this class.

I am somewhat familiar _____ with this class.

I am not too familiar _____ with this class.

Signed _____

Date _____

SPONSOR IMPLEMENTATION MEASURE

MODEL: HIGH/SCOPE

SITE:

KEY: 1 = Specification not at all implemented.
2 = Specification implemented to some extent.
3 = Specification implemented to a great extent.
4 = Technically implemented.

You are rating the class with the following personnel, please circle, beneath each teacher's name, the year in which he or she is now teaching in the model.

Teacher 1: 1st 2nd 3rd	Teacher 2: 1st 2nd 3rd	Teacher 3: 1st 2nd 3rd
1. The organization of the room and the equipment, both specific and general items are as the model specifies.		Classroom: 1 2 3 4
2. The routines are consistent and their content is as the model suggests.		Classroom: 1 2 3 4
3. The teachers use divergent questions and in general are stimulating the use of language in the classroom.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
4. Teachers understand and are correctly sequencing activities, starting from the index level through all levels of representation.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
5. Teachers understand and appropriately present to children the concepts of temporal relations, spatial relations, classification, seriation.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
6. Teachers give time every day to planning and evaluating.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
7. Teachers visit parents regularly during the year, implementing the purposes of home-school contact.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4

HIGH SCOPE

SITE:

8. The teachers work with volunteers is according to model specifications.

Teacher 1: 1 2 3 4

Teacher 2: 1 2 3 4

Teacher 3: 1 2 3 4

I am very familiar _____ with this class.

I am somewhat familiar _____ with this class.

I am not too familiar _____ with this class.

Signed _____

Date _____

SPONSOR IMPLEMENTATION MEASURE

MODEL: FLORIDA

SITE:

KEY: 1 = Specification not at all implemented.
2 = Specification implemented to some extent.
3 = Specification implemented to a great extent.
4 = Technically implemented.

You are rating the class with the following personnel, please circle, beneath each teacher's name, the year in which he or she is now teaching in the model.

_____

Parent Educator: 1st 2nd 3rd

1. The PAC is involved in the operation of the model, hiring P.E.'s and teachers and making decisions regarding curriculum and tasks.	PAC:	1 2 3 4
2. The P.E. is making weekly visits and is working with the teacher in planning and evaluating these visits.	P.E.:	1 2 3 4
3. The P.E. gives tasks to her group of M.O.'s with full understanding of the purposes of the tasks and appropriately uses a range of alternatives in presenting the task to M.O.'s.	P.E.:	1 2 3 4
4. The P.E. employs the seven desirable teaching behaviors in her visits with the M.O.'s.	P.E.:	1 2 3 4
5. The P.E. works with M.O.'s in such a way that they understand the tasks, are involved with the whole task process, as evidenced by their talking a great deal, enjoying the task, contributing to expansion of the given task, as well as suggesting new tasks.	P.E.:	1 2 3 4

I am very familiar ____ with this Parent Educator.

I am somewhat familiar ____ with this Parent Educator.

I am not too familiar ____ with this Parent Educator.

Signed _____

Date _____

SPONSOR IMPLEMENTATION MEASURE

MODEL: EDC

SITE:

KEY: 1 = Characteristic not at all implemented.
2 = Characteristic implemented to some extent.
3 = Characteristic implemented to a great extent.
4 = Technically implemented.

You are rating the class with the following personnel, please circle, beneath each teacher's name, the year in which he or she is now teaching in the model.

Teacher 1: 1st 2nd 3rd	Teacher 2: 1st 2nd 3rd	Teacher 3: 1st 2nd 3rd
1. The classroom is arranged so that it is functional, unique, and enables children to take responsibility for their own actions and learning.		Classroom: 1 2 3 4
2. A variety of materials are provided, and children are allowed to use them in unprescribed ways.		Classroom: 1 2 3 4
3. Activities in the program are balanced, the teaching goals are incorporated into routine activities.		Classroom: 1 2 3 4
4. Planning time for teachers and aides is used; the aide both teaches and shares housekeeping chores; classroom volunteers are oriented to the model.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
5. Teachers have confortable relations with the children and are sensitive to their needs; they encourage the child's best efforts in a way most appropriate for each child; they value children's ideas.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
6. Teachers encourage independence, imagination and persistence in chosen tasks.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
7. Children take initiative and responsibility for their own work, which includes knowing how to get help when they need it, refusing help when appropriate, being intensely involved, using materials and activities in a variety of ways, and learning from their mistakes.		Children: 1 2 3 4

EDC PAGE TWO

SITE:

8. While the children take an interest in each other -- by helping and listening to each other, they are also free to argue things out both among themselves and the teacher, and there is a place in the class for a child to be by himself.	Children:	1 2 3 4
9. The children feel good about themselves, demonstrate curiosity and a variety of opinions, and show signs of humor, fun, and joyousness.	Children:	1 2 3 4
10. There is an ongoing effort to help parents understand the model and to report to the parents about their children.	Parents:	1 2 3 4

I am very familiar _____ with this class.

I am somewhat familiar _____ with this class.

I am not too familiar _____ with this class.

Signed _____

Date _____

SPONSOR IMPLEMENTATION MEASURE

MODEL: PITTSBURGH

SITE:

KEY: 1 = Specification not at all implemented.
2 = Specification implemented to some extent.
3 = Specification implemented to a great extent.
4 = Technically implemented.

You are rating the class with the following personnel, please circle, beneath each teacher's name, the year in which he or she is now teaching in the model.

Teacher 1: 1st 2nd 3rd	Teacher 2: 1st 2nd 3rd	Teacher 3: 1st 2nd 3rd
1. There are two main components to the schedule: a prescribed learning activity period and an exploratory learning activity period; usually the first period is from 30-45 minutes long and exploratory learning ends around the second hour of the day. Group time is an informal general time for conversation between members of the class and teachers.		Classroom: 1 2 3 4
2. Materials are organized the room according to various learning centers; teaching materials for prescribed learning are keyed to the objectives included in each of the curriculum and are clearly labeled so that children can find them.		Classroom: 1 2 3 4
3. The staff consists of one teacher and one teacher aide. Staff members rotate responsibilities of a) testing and tutoring, and b) "traveling," and they jointly develop tasks for the children.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
4. Teachers assign tasks for every curriculum component of the program, gratification, classification and perceptual development, they promote communication skills and encourage independent learning.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4
5. In order to avoid giving the child a feeling of failure, when assigning tasks, teachers provide different ways for a child to master a task if it is too difficult for him, and they treat diagnostic tasks as guides for learning activities.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4

PITTSBURGH

PAGE TWO

SITE:

6. Teachers reinforce children for completion of tasks and give verbal feedback on tasks; teachers do not give attention reinforcement to children who leave tasks but make subtle but neutral attempts to encourage them to complete tasks.	Teacher 1:	1 2 3 4
	Teacher 2:	1 2 3 4
	Teacher 3:	1 2 3 4
7. Children are confident of their ability to learn and can cope with their surroundings.	Children:	1 2 3 4
8. Children move freely about the environment and choose their own activities during the exploratory period; children are experiencing success.	Children:	1 2 3 4

I am very familiar _____ with this class.

I am somewhat familiar _____ with this class.

I am not too familiar _____ with this class.

Signed _____

Date _____

SPONSOR IMPLEMENTATION MEASURE

MODEL: REC

SITE:

KEY: 1 = Specification not at all implemented.
2 = Specification implemented to some extent.
3 = Specification implemented to a great extent.
4 = Technically implemented.

You are rating the class with the following personnel, please circle, beneath each teacher's name, the year in which he or she is now teaching in the model.

Teacher 1: 1st 2nd 3rd	Teacher 2: 1st 2nd 3rd	Teacher 3: 1st 2nd 3rd
1. The room is clearly defined into functional learning areas to stimulate the children to explore on their own; children are able to flow from one area to another; they learn through play and the use of manipulative materials; they usually work alone or in spontaneously formed groups.		Classroom: 1 2 3 4
2. Classroom materials which meet one or more of the model's criteria are attractively arranged and within easy reach of children; the materials are varied occasionally to stimulate exploration; children choose activities and materials freely 25-50% of the time.		Classroom: 1 2 3 4
3. Basic activities are language arts, construction and building, art, math, and housekeeping; the schedule is flexible and not all activities occur each day; there is a special time for math in three or four small groups with an adult (number varying with number of volunteers).		Classroom: 1 2 3 4
4. Each class has one teacher and one aide working with the children. Parent volunteers are optional but desirable.		Classroom: 1 2 3 4
5. The adults work with the children in using the Talking Page by introducing it to small groups during activity time; following up with children who choose the Talking Page lesson as an activity; and going over the Talking Page progress check with the children individually.		Teacher 1: 1 2 3 4 Teacher 2: 1 2 3 4 Teacher 3: 1 2 3 4

REC

PAGE TWO

SITE:

6. Adults are sensitive to individual needs: they encourage children to help themselves; they provide for children to experience success at their own pace; the teacher provides an individualized program and joins the children as frequently as possible in their self chosen activities in order to fulfill her role as a catalyst for the next logical step in the learning process.	Teacher 1:	1 2 3 4
	Teacher 2:	1 2 3 4
	Teacher 3:	1 2 3 4
7. Adults elicit as much language as possible from the children.	Teacher 1:	1 2 3 4
	Teacher 2:	1 2 3 4
	Teacher 3:	1 2 3 4
8. Children seem to know what to do in the classroom: they work freely and independently; they know how to use adults properly; when they have a problem, they solve it themselves; and they actively seek and select what they will do.	Children:	1 2 3 4
9. Children show evidence of a positive self image.	Children:	1 2 3 4
10. Disruptive behavior seldom occurs in the classroom, and the children seldom argue or fight.	Children:	1 2 3 4

I am very familiar _____ with this class.

I am somewhat familiar _____ with this class.

I am not too familiar _____ with this class.

Signed _____

Date _____

00096

3. Sponsor Overall Ratings

Head Start Planned Variation 1971-1972

Sponsor Implementation Measure

Model _____

Please give an overall rating to the implementation of each of the classes listed below using the following definitions:

1 = not implementing the model
2 = implementing the model to some extent
3 = implementing the model to a great extent
4 = completely implementing the model

Site: _____

Class	Head Teacher				
1	_____	1	2	3	4
2	_____	1	2	3	4
3	_____	1	2	3	4
4	_____	1	2	3	4
5	_____	1	2	3	4
6	_____	1	2	3	4

Site: _____

Class	Head Teacher				
1	_____	1	2	3	4
2	_____	1	2	3	4
3	_____	1	2	3	4
4	_____	1	2	3	4
5	_____	1	2	3	4
6	_____	1	2	3	4

Site: _____

Class	Head Teacher				
1	_____	1	2	3	4
2	_____	1	2	3	4
3	_____	1	2	3	4
4	_____	1	2	3	4
5	_____	1	2	3	4
6	_____	1	2	3	4

Site: _____

Class	Head Teacher				
1	_____	1	2	3	4
2	_____	1	2	3	4
3	_____	1	2	3	4
4	_____	1	2	3	4
5	_____	1	2	3	4
6	_____	1	2	3	4